"Issues in a Regulated Insurance Market: The Case of Health Care"

Dissertation
for the Faculty of Economics, Business Administration and Information Technology of the University of Zurich

to achieve the title of
Doctor of Philosophy
in Economics

presented by
Michèle Sennhauser
from CH-Lütisburg SG
& CZ-Brno JM

approved in April 2011 at the request of
Prof. Dr. Peter Zweifel
Prof. Dr. Konstantin Beck

The Faculty of Economics, Business Administration and Information Technology of the University of Zurich hereby authorises the printing of this Doctoral Thesis, without thereby giving any opinion on the views contained therein.

Zurich, 06.04.2011

Chairman of the Doctoral Committee: Prof. Dr. Dieter Pfaff

Preface

This doctoral thesis was written during my time as a research assistant at the Department of Economics (formerly the Socioeconomic Institute) of the University of Zurich, Switzerland, from 2006 to 2011. I enjoyed my time there a lot and learned an incredible amount, thanks to the following people.

I would first like to thank my supervisor Prof. Dr. Peter Zweifel for his invaluable support and guidance. I am grateful for his encouragement and challenges during my master's studies, for his constructive intellectual input and his advice about the essays in this dissertation, and for his continuing support in further education. He was also generous enough to give me the opportunity to study at the University of Chicago. I would also like to thank him for giving me responsibility for several interesting empirical projects. I not only learned a lot from investigating these projects, but I also enjoyed them a lot. I am indebted to Prof. Dr. Konstantin Beck for evaluating and grading this thesis. During my master's studies, he introduced me to the issue of risk adjustment in health insurance and awoke my interest in the problems of the Swiss health care system. Moreover, I thank him for providing data for the simulations in Chapters 4 and 5. As previously mentioned, I have spent one year of my doctoral studies at the University of Chicago. Thanks to Prof. Willard G. Manning, I benefited tremendously during this time. I would like to thank him for sharing his time, his patience, his econometric advice, and his deep knowledge of econometrics.

I would also like to thank my colleagues and co-workers. First, I want to mention Dr. Hyo Jung Tak. During my stay in Chicago, she shared her extensive specialist

knowledge and, more importantly, her friendship with me. I am indebted to my co-authors Dr. Patrick Eugster and Dr. Johannes Schoder, as well as to my fellow workers Angelika Brändle, Dr. Boris Krey, Philipp Morf, Dr. Ilja Neustadt, Maurus Rischatsch, Adrian Rohner, Micha Ruflin, PD Dr. Felix Schläpfer, Dr. Harry Telser, Maria Trottmann, and Philippe Widmer. It will always be a great pleasure to remember our lively discussions across a wide range of topics. Additionally, I want to thank the whole team at the library accommodating my numerous literature requests. I very much appreciated their reliable, unbureaucratic, and professional services.

I would also thank all the other people who supported and accompanied me through my education.

This is also to my parents Michaela and Marcel Sennhauser, who directly supported me during my master's studies and provided invaluable support during my doctoral studies. I also want to thank my grandparents Pia and Josef Sennhauser, who always believed in me without any doubts.

Most of all, I want to thank Simon J. Bolt. During the last few years, he has offered me invaluable support and sustained me through this journey. I thank him for his advice, his patience, and his critical mind. Thank you!!

Michèle Sennhauser
Zurich, December 2010

Vorwort

Diese Dissertation entstand während meiner Tätigkeit als Forschungsassistentin am Institut für Volkswirtschaftslehre (ehemals Sozialökonomisches Institut) der Universität Zürich von 2006 bis 2011. Ich habe diese Zeit sehr genossen. Dies habe ich den folgenden Menschen zu verdanken.

Zuerst möchte ich mich bei Prof. Dr. Peter Zweifel für seine Unterstützung und für das in mich gesetzte Vertrauen bedanken. Ich danke ihm für seinen fordernden und fördernden Unterrichtsstil während dem Studium, für seine hilfreichen und wertvollen Denkanstösse zu den Essays dieser Dissertation, für seine Unterstützung bei Weiterbildungen und dass er mir die Möglichkeit gegeben hat an der University of Chicago zu studieren. Ich danke ihm auch dafür, dass er mir die Verantwortung für viele spannende empirische Projekte übertragen hat. Dabei habe ich sehr viel gelernt und die Arbeit hat mir grossen Spass bereitet. Mein Dank gilt weiter Prof. Dr. Konstantin Beck für die Übernahme des Koreferats. Bereits während dem Studium hat er mich für das Thema des Risikoausgleichs zwischen den Krankenversicherern und die verschiedenen Probleme des schweizerischen Gesundheitswesens begeistert. Ich danke ihm auch für die Bereitstellung der Daten, die den Berechnungen der Kapitel 4 und 5 zu Grunde liegen. Wie bereits erwähnt, durfte ich ein Jahr meines Doktorandenstudiums an der University of Chicago verbringen. Dass ich in dieser Zeit so viel gelernt habe, verdanke ich Prof. Willard G. Manning. Er hat sich sehr viel Zeit genommen um mir ökonometrische Ratschläge zu geben und hat mein Verständnis der Materie entscheidend verbessert.

Ich möchte mich auch herzlich bei all meinen Kolleginnen und Kollegen bedanken. Zuerst möchte ich Dr. Hyo Jung Tak erwähnen. Sie half mir nicht nur bei fachlichen Fragen, sondern hat dazu beigetragen, dass Chicago auch ausserhalb der Uni Spass machte. Weiter möchte ich meinen Kollegen aus Zürich, beginnend bei meinen beiden Koautoren Dr. Patrick Eugster und Dr. Johannes Schoder, aber auch Angelika Brändle, Dr. Boris Krey, Philipp Morf, Dr. Ilja Neustadt, Maurus Rischatsch, Adrian Rohner, Micha Ruflin, PD Dr. Felix Schläpfer, Dr. Harry Telser, Maria Trottmann und Philippe Widmer herzlich danken. Die vielen Diskussionen verschiedenster Natur werden mir immer in sehr guter Erinnerung bleiben. Dem ganzen Bibliotheksteam danke ich für die sehr zuverlässige und speditive Beschaffung meiner zahlreichen Literaturwünsche. Ich schätzte die unkomplizierte und unbürokratische Art sehr.

Ich möchte mich auch bei allen anderen Personen bedanken die mich in irrgend einer Form während meiner Ausbildung begleitet und unterstützt haben.

Einen speziellen Dank möchte ich meinen Eltern Michaela und Marcel Sennhauser aussprechen. Durch ihre Unterstützung haben sie mir das Studium direkt, und das Doktorandenstudium indirekt sehr erleichtert. Ich möchte mich auch bei meinen Grosseltern Pia und Josef Sennhauser bedanken die immer ohne jeglichen Zweifel an mich geglaubt haben.

Zu guter Letzt danke ich Simon J. Bolt. Er war für mich während dieser Zeit meine grösste Stütze und stand mir in jeder Situation zur Seite. Ich danke ihm für seine Ratschläge, seine Geduld und für die vielen kritischen Diskussionen. Vielen herzlichen Dank!!!

Michèle Sennhauser
Zürich, Dezember 2010

Contents

1 Introduction **3**

2 A New Pharmaceutical: Is it Worth the Money? Whose Money? **11**

2.1 Introduction 11

2.2 Literature Review 14

2.3 The Experiment 16

2.4 Ex-ante vs. Ex-post Willingness-To-Pay 22

2.5 Hypotheses 26

2.6 Data: Descriptive Statistics 28

2.7 Empirical Results 31

2.7.1 Willingness-To-Pay 31

2.7.2 Willingness-To-Pay across Subgroups 35

2.8 Conclusions 37

3 Why the Linear Utility Function is a Risky Choice in Discrete-Choice Experiments **43**

3.1 Introduction 43

3.2 Theoretical Foundations 45

3.2.1 Discrete-Choice Experiments 45

3.2.2 Specification of the Utility Function 47

3.2.3 Assessing Goodness-of-Fit 49

3.3 The Experiment 51

3.3.1 Background: Swiss Statutory Social Health Insurance 51

3.3.2 Attributes 52

3.3.3 Pretest and Design . . . 53
3.3.4 Sample and Interview Strategy . . . 55
3.4 Empirical Results . . . 56
3.4.1 Specification of the Utility Function . . . 56
3.4.2 Comparison of Goodness-of-Fit . . . 60
3.4.3 Comparison of Willingness-to-Pay . . . 63
3.5 Conclusions . . . 68

4 Capping Risk Adjustment? 73
4.1 Introduction . . . 73
4.2 Data Basis, Descriptive Statistics, and Representativeness . . . 75
4.3 Hospitalization as an Additional Criterion . . . 79
4.4 Limiting the Volume of Risk Adjustment . . . 81
4.5 Optimizing the Cap on the Volume of Risk Adjustment . . . 85
4.6 Consequences of Capping Risk Adjustment . . . 90
4.6.1 Theoretical Considerations . . . 90
4.6.2 Empirical Illustration . . . 92
4.7 Conclusions . . . 95

5 Fine-Tuning of Health Insurance Regulation: Unhealthy Consequences for an Individual Insurer 101
5.1 Introduction . . . 101
5.2 Simulation of Risk Adjustment Values and Data Basis . . . 104
5.2.1 Methodology . . . 104
5.2.2 Data Basis . . . 105
5.3 Simulating the Impacts of the New RA Formula . . . 109
5.3.1 Risk Adjustment with the New Criterion . . . 109
5.3.2 Impacts on Risk Adjustment Payments by Health Insurer A . . . 110
5.3.3 Impact on Risk Management . . . 114
5.4 Conclusions . . . 116

6 Conclusions 121

List of Figures

2.1 Choice question example: Fixed status quo (current insulin) vs. alternative (new insulin) 22

3.1 Choice question example: Fixed status quo (current contract) vs. alternative (new contract) 54
3.2 Smoothed scatter plots for *Deductible* and *Contribution* 57
3.3 Plots of estimated probit regression coefficients vs. approximate quartile midpoints of *Deductible* and *Contribution* 58
3.4 Regression result modified Hosmer-Lemeshow test linear utility function 62
3.5 Regression result modified Hosmer-Lemeshow test non-linear utility function 62
3.6 95 percent confidence ellipses for marginal WTP 67

4.1 Cross-subsidies by age group, persons without / with hospitalization or living in nursing home during the previous year, canton of Zurich, CHF (2005) 81
4.2 HCE of women by age, actual, estimated (three insurers), and official, Switzerland, CHF (2005) 92

5.1 Official RA values according to age and gender, canton of Zurich, CHF (2005) 108
5.2 Estimated RA values with and without hospitalization according to age and gender, canton of Zurich, CHF (2005) 108
5.3 Hospitalization rate, insurer A vs. simulated nationwide values, men (2005) 112
5.4 Inpatient cost, insurer A vs. simulated nationwide values, CHF (2005) . 113

5.5 Outpatient cost, insurer A vs. simulated nationwide values, CHF (2005) 113

List of Tables

2.1 DCE attributes, labels and levels 19
2.2 Descriptive statistics 30
2.3 Results of a random-effects probit estimation, aggregate sample 31
2.4 Marginal WTP for product attributes, aggregate sample, Euro 33
2.5 WTP for product attributes, aggregate sample, Euro 34
2.6 WTP for product attributes, stratified by diabetes type 36

3.1 DCE attributes, labels and levels 52
3.2 Selected descriptive statistics 55
3.3 Interactions and terms of higher order resulting from Step 4 59
3.4 Comparison of goodness-of-fit results linear and non-linear utility function 60
3.5 Marginal WTP linear and non-linear utility function, CHF per month . 65
3.6 Results random-effects probit estimation, linear and non-linear utility function 70

4.1 Volume of cross-subsidy per canton, CHF (2005) 77
4.2 Simulated and official cross-subsidies per capita according to age and gender, CHF (2005) 78
4.3 Cross-subsidization without and with the hospitalization adjuster, CHF (2005) 80
4.4 Volumes of cross-subsidization (CS) and risk adjustment (RA), CHF (1996-2005) 82
4.5 Capping the volume of cross-subsidies, canton of Zurich, CHF (2005) . 93
4.6 Change of CS and RA volumes, CHF bn. (2005) 95

Chapter I
Introduction

1 Introduction

This dissertation consists of four essays dealing with issues in the regulated insurance market of health care. Contemporary political discussions about health care mostly focus on rising health care expenditure. In the hope of curbing this surge, several countries have instituted competition between social health insurers and introduced the concept of "managed competition" (Van de Ven et al., 2007). However, social health insurers are still subject to many regulations (Enthoven, 1986), including the imposition of risk-independent premiums. In the United States, these premiums are community-rated at the level of a firm's employees; in the Netherlands and Germany, the premium comes from a uniform contribution rate derived from labor income; and in Switzerland, uniform premiums per capita are charged by a given insurer (Zweifel and Breuer, 2006). Due to this regulation, health insurers have incentive to select favorable risks (Pauly, 1984). Enrolling high risks leads to expected losses, while enrolling low risks leads to expected gains. However, for the economy as a whole, risk selection is a zero-sum game and yields higher health care expenditure across the board.

According to Zweifel et al. (2009, Chapter 7), compulsory health insurance and risk-independent premiums allow three types of measures to avoid risk selection: (1) An open enrollment policy that prevents direct risk selection (i.e., an insurer controls contract signatories based on observable characteristics correlated to risk type); (2) Regulation of the benefit package that prevents indirect risk selection (i.e., an insurer influences enrollment indirectly by offering contracts with benefits that appeal to certain risk types); and (3) Risk adjustment to reduce incentives for risk selection. Risk adjustment makes insurers with an above-average share of favorable risks cross-subsidize insurers with many unfavorable risks. While an open enrollment

policy underlies all chapters of this dissertation, Chapters 2 and 3 deal specifically with measure (2), while Chapters 4 and 5 contribute to the topic of measure (3).

Chapters 2 and 3 provide experimental measurements of preferences for the list of benefits. From a theoretical viewpoint, minimum as well as maximum benefits should be defined, forcing insurers to offer important high-risk benefits in addition to preventing them from including services that particularly appeal to low risks. However, decisions about inclusion of a specific treatment or pharmaceutical product on the list of benefits should be made according to the insured's willingness-to-pay and the costs of the benefit. Chapter 2 presents evidence for the case of a new pharmaceutical product, viz. a long-acting insulin analogue. A discrete-choice experiment is conducted in Germany to estimate willingness-to-pay of social statutory health insurance (GKV) members, including both actual patients (type 1 and insulin-dependent type 2 diabetics) and potential patients (non-diabetics and insulin-naive type 2 diabetics). By including two financial attributes (copayment and health insurance contributions) preferences for the mode of financing were elicited.

Compared to Germany, Switzerland offers more elements of choice in statutory health insurance. These elements give health insurers more opportunities for risk selection, but consumers also benefit from this diversity, being able to choose a contract that corresponds more closely to their preferences. Chapter 3 reexamines a discrete-choice experiment conducted by Becker (2006, Chapters 6-8) in Switzerland, estimating willingness-to-pay of the Swiss population for various proposed reforms concerning the list of benefits (reimbursement of alternative medicine treatments, restriction to generics, and immediate access to new treatment methods), as well as financing (copayment and deductible levels).

Chapters 4 and 5 focus on risk adjustment as the third option of secondary regulation to prevent risk selection. The design of an optimal risk-adjustment formula is a widely discussed topic. Since its introduction in 1996, the Swiss formula has been based solely

on age and gender. Effective in 2012, the formula will include a third risk adjuster, "Hospitalization or living in a nursing home more than three days during the previous year", to mitigate risk selection. These two chapters analyze the consequences of this refinement. Chapter 4 has two objectives. First, it analyzes the effects of the third risk adjuster on the volume of risk adjustment (or, more precisely, on the volume of cross-subsidization between the insured). Second, it addresses how a cap on the volume of risk adjustment can be placed so that opportunity costs in terms of incentives for risk selection are minimized.

A case study forms the core of Chapter 5. This study shows the consequences of this fine-tuning on a health insurer that seems not to be a "cherry-picker", but an insurer among the forefront in conforming to the stated objectives of Swiss health policy, i.e. to achieve savings through Managed Care.

The essays adopt different points of view. Chapters 2 and 3 consider the insured's preferences, with distinctions of high and low risks central to Chapter 2. Consequences of health insurance regulation for the economy as a whole are presented in Chapter 4, while Chapter 5 focuses on a single health insurer's perspective. However, all the essays are based on individual panel data. More than 1,100 enrollees of the German statutory health insurance were interviewed face-to-face for the discrete-choice experiment of Chapter 2. For the experiment presented in Chapter 3, 1,000 telephone interviews of enrollees of Swiss statutory health insurance were conducted. Three large Swiss health insurers provided data about holders of basic health insurance over five years for Chapters 4 and 5, resulting in records of 2.78 million insured. This number corresponds to an average market share of more than 25 percent. Additionally, a medium-sized health insurer provided data for the simulations of Chapter 5. Concerning methodological common grounds, Chapters 2 and 3 share the discrete-choice approach. Econometric estimation of health care expenditure underlies the results of Chapters 4 and 5.

The topic of this dissertation can also be considered from an efficiency point of view. If the list of benefits in statutory social health insurance includes treatments or pharmaceuticals that the insured exhibit a willingness-to-pay lower than costs, health care resources are used inefficiently. In the case of Chapter 2, insulin-dependent diabetics might have a high enough (ex-post) willingness-to-pay for the new insulin product. However, for the reimbursement to be justified, ex-ante willingness-to-pay by the non-affected insured has to be sufficiently high as well.

Using the results of discrete-choice experiments as the basis for decision-making may lead to efficiency if estimated willingness-to-pay indeed corresponds to actual willingness-to-pay. Recent literature shows many problems and challenges, such as designing questionnaires and applying the newest econometric methods (Hoyos, 2010). Chapter 3 contributes to this literature by showing that the commonly assumed linear form of the utility function may be a risky choice in discrete-choice experiments.

Risk adjustment is generally introduced to mitigate incentives for risk selection. However, as explored in Chapter 4, this can also have effects on insurers' quest for efficiency. A refinement of the risk-adjustment formula would increasingly shelter insurers from financial risk, undermining their incentive to improve efficiency. For this reason, a government might consider simply capping the volume of risk adjustment. This then exposes insurers to some residual financial risk. Chapter 4 further studies how risk-adjustment values can be reduced to minimize the increase in health care expenditure variance borne by the insurer, thus also minimizing the risk-selection effort.

Insolvency and the eventual market exit of insurers who only survived thanks to cream skimming may be considered to be efficiency enhancing. However, Chapter 5 outlines how the fine-tuning of the risk-adjustment formula may also cause an innovative health insurer, who had successfully implemented Managed Care to lower rates of hospitalization, to go bankrupt. This threat to survival may trigger adjustments in the insurer's risk-management strategy, causing efficiency losses to tax-payers, employees,

patients, and the economy as a whole.

Chapter 6 concludes by stating the major policy implications and disclosing possible future extensions.

Note that Peter Zweifel co-authored Chapters 2, 4, and 5, Patrick Eugster co-authored Chapter 4, and Johannes Schoder co-authored Chapter 5. I am indebted to Willard G. Manning for the intellectual input of Chapter 3. Chapter 2 is submitted to *PharmacoEconomics*, and Chapter 3 to *Health Economics*. Chapter 4 is published in the *Journal of Health Economics* and Chapter 5 in the *International Journal of the Economics of Business*.

Michèle Sennhauser
Zurich, December 2010

Chapter II
A New Pharmaceutical: Is it Worth the Money? Whose Money?

Michèle Sennhauser and Peter Zweifel

Submitted to

PharmacoEconomics

2 A New Pharmaceutical: Is it Worth the Money? Whose Money?

2.1 Introduction

Health care expenditure (HCE) and especially pharmaceutical expenditure is rising in almost all developed countries. For example, in the United States the share of pharmaceutical expenditures in total HCE increased from 9 percent in 1996 to 12 percent in 2008 (OECD, 2010). In an attempt to curb this surge, several countries have introduced a cost-effectiveness standard for new pharmaceuticals. This led to the creation of the Medicare Payment Advisory Commission (MEOPAC) scheme in Australia, the National Institute for Clinical Excellence (NICE) in the United Kingdom, and the Institute for Quality and Efficiency in Health Care (IQWiG) in Germany. In Germany, the pharmaceutical bill paid for by statutory health insurance (GKV) increased from Euro 22 billion (bn.) in 2004 to Euro 26 bn. in 2007, or from 1.00 percent of GDP to 1.07 percent (Statistical Offices of the Länder, 2009). Before 2007, pharmaceutical innovations had to meet safety and efficacy benchmarks to be included in the GKV list of benefits. Now, they also have to be cost-effective.

This study seeks to provide evidence for deciding whether or not a new pharmaceutical for insulin therapy, a long-acting insulin analogue[1], should be included in

[1]The product considered in this chapter is "Insulin detemir" by Novo Nordisk Pharma GmbH. Modern insulin therapy uses long- and short-acting insulin in combination. Whereas rapid-acting insulin meets insulin need during mealtimes, long-acting insulin assures base-level supply. Both rapid- and long-acting insulin can be human or insulin analogue. Whereas human insulin is genetically identical to insulin from the human pancreas, insulin analogue differs slightly to improve the insulin's properties.

the German benefit list of social health insurance. So far, the standard of treatment has been neutral protamine Hagedorn (NPH) insulin which is human insulin. The new pharmaceutical promises several medical advantages, such as fewer events of hypoglycemia, less weight gain (or even weight loss), easier preparation, and more flexibility in injection time (for a list of references on clinical outcomes studies, see Section 2.3 below). These potential advantages come with an average cost of Euro 226 per year and diabetic (in Germany). Concerning the cost-effectiveness of insulin analogues compared to NPH insulin, there have been several studies presenting mixed, but mostly positive results. Whereas e.g. Cameron and Bennett (2009) find the pharmaceutical not to be cost-effective, other studies disagree, e.g. Valentine et al. (2007) for type 2 diabetics[2] in the United States.

There are two reasons why this preparation is of special interest. First, diabetes prevalence is higher then ever in industrialized countries and continues to increase rapidly. The World Health Organization projects the number of diabetics worldwide to rise from 170 million (mn.) in 2000 to 360 mn. by 2030 (World Health Organization, 2007, Wild et al., 2004). For the United States Huang et al. (2009) estimate the number of diabetics to increase from 23.7 mn. in 2009 to 44.1 mn. patients in 2034. Annual diabetes-related expenditure is expected to rise from USD 113 bn. to USD 336 bn. during the same period. The prevalence of diabetes in Germany is 4 to 10 percent between ages 40 and 59 and 18 to 28 percent for ages above 60 (Hauner, 2008). Second, long-acting insulin analogues may well constitute a test case. IQWiG recommended to drop long-acting insulin analogues from the benefit list, judging it not to be cost-effective (IQWiG, 2009 and IQWiG, 2010). However, these recommendations did not take into account preferences of (potential) patients. Several aspects of the drug which may be innovative from the patient's perspective were neglected or judged

[2]In case of type 1 diabetes the body does not produce insulin. It is usually diagnosed in children and young adults and has to be treated with insulin from the beginning. In type 2 diabetes, either the body does not produce enough insulin or the cells ignore the insulin. This type is usually diagnosed in the elderly. Diabetics of type 2 are called "insulin-naive" if they are not treated with insulin (yet) but with oral anti-diabetics. However, during the course of their disease they will need insulin treatment as well (American Diabetes Association, 2010).

as therapeutically unimportant. The (potential) patients' preferences can be elicited in a discrete-choice experiment (DCE). With the inclusion of a financial attribute, willingness-to-pay (WTP, or willingness-to-accept (WTA)) values can be attached to the characteristics of the product, permitting to express its (dis-)utility in terms of money. From the point of view of the insured (comprising both actual and potential patients), inclusion of the new product in the list of benefits is justified if they exhibit a WTP that exceeds the extra cost of the treatment.

To the knowledge of the authors, there has been no WTP study concerning long-acting insulin analogues. This study presents a DCE comparing insulin analogue with NPH insulin conducted in Germany in the Fall of 2007. Participants in the DCE are 1,110 members of statutory health insurance GKV, of whom 200 suffer from type 1 diabetes, 150 from insulin-treated type 2 diabetes, and 150 from insulin-naive type 2 diabetes. Distinguishing these groups allows to estimate ex-ante WTP for non-diabetics and insulin-naive diabetics on the one hand and ex-post WTP for insulin-treated patients on the other. Four attributes describing differences in insulin therapy between NPH insulin and insulin analogue were included according to medical outcomes studies: Risk of hypoglycemia, weight gain during the first six months of insulin treatment, need to swing (not shake) the insulin before injections, and flexibility with regard to time of injection. There are two attributes for the mode of payment, financing through patients themselves (copayment) and through increased health insurance contributions, respectively. The inclusion of two financial attributes permits to test whether the new drug has a favorable benefit-cost ratio regardless of the boosting of WTP caused by health insurance.

There are four main questions to be answered. (1) Is there positive WTP for long-acting insulin analogue by the members of German statutory health insurance? (2) If so, which product attributes contribute to WTP? (3) Is there preference heterogeneity between non-affected non-diabetics and insulin-naive type 2 diabetics on the one hand and type 1 diabetics and insulin-treated type 2 diabetics on the other? (4) Is the

benefit-cost ratio of the new drug favorable regardless of whether it is financed jointly through increased GKV contributions or by patients through copayment?

This chapter is organized as follows. Section 2.2 gives an overview of cost-effectiveness studies concerning insulin analogue and of preference studies regarding insulin therapy. Section 2.3 presents the interview strategy and questionnaire design with the attributes and levels. Then theory behind DCEs is briefly presented in Section 2.4 with emphasis on the difference between ex-ante and ex-post WTP measurement. Hypotheses are formulated in Section 2.5 before presenting descriptive statistics in Section 2.6. Section 2.7 contains the empirical evidence and hypothesis tests. The four questions raised are answered in the concluding Section 2.8.

2.2 Literature Review

Cost-Effectiveness Studies

Existing cost-effectiveness studies of the insulin analogue use quality-adjusted life years (QALYs) as the benefit measure and the incremental cost effectiveness ratio (ICER) as the valuation criteria. Until recently, they focused on the treatment of type 1 patients. For the UK, Palmer et al. (2004) and Palmer et al. (2007) find improvements of 0.09 and 0.66 QALYs, resulting in ICER of GBP 19,285 and GBP 2,500, respectively, which compare favorable with the ICER of GBP 30,000 used by NICE. These estimates are confirmed by Palmer et al. (2008) for Denmark with an ICER of DKK 55,867 or GBP 6,600. In their multi-country study, Gschwend et al. (2009) conclude that the insulin analogue is likely to be a dominant treatment strategy for type 1 patients in Belgium, Germany, and Spain, and highly cost-effective in France and Italy with an ICER of Euro 519 and Euro 3,256 per QALY, respectively. For the United States, Leichter (2008) found the pharmaceutical to be cost-effective due to lower incidence of acute hypoglycemic events and costly, chronic complications such as nephropathy. In the same vein, Valentine et al. (2006) estimate an ICER of USD 14,974. With regard to type 2 patients the findings are slightly more mixed.

While Valentine et al. (2007) estimate an even lower ICER of USD 6,269 than for type 1 patients, Cameron and Bennett (2009) arrive at USD 387,729, leading them to conclude that long-acting analogues are unlikely to present an efficient use of health care resources.

WTP Studies

For all its popularity, the cost-effectiveness measure is not satisfactory from an economic point of view for two main reasons. First, QALYs focus exclusively on health outcomes, neglecting attributes of the treatment process such as fear, isolation, and confinement. Second, this measure does not allow to pit resources devoted to health against resources devoted to other uses. Specifically, it fails to reflect the preferences of citizens who may favor an expansion of the health budget, with the consequence that the threshold ICER value (e.g. the GBP 30,000/QALY applied by NICE) could be adjusted upward. By way of contrast, measurement of WTP values permits to compare marginal benefit to marginal cost, both expressed in money.

The first WTP study concerning insulin therapy is Davey et al. (1998) in Australia. The authors compared insulin lispro, the first rapid-acting insulin analogue, with neutral (regular) insulin using a contingent-valuation approach. Respondents first were presented with the descriptions of two types of insulins and had to choose one. Then, they were taken through a series of "bid-up" questions to determine their maximum WTP. The sample consisted of both type 1 and type 2 diabetics who had been treated with insulin before. The same method was applied by Dranitsaris et al. (2000) to elicit WTP for the rapid-acting insulin analogue Humalog Mix 25. Unlike the first study, the sample was drawn from the general tax-paying public. Sadri et al. (2005) analyzed WTP for inhaled insulin, using the payment scale method. The study involved type 1 and type 2 diabetics and presented results both for insulin-naive and insulin-dependent patients.

In contrast to the contingent-valuation approach, the levels of all attributes characterizing the alternative are allowed to change in a DCE, which makes participants repeatedly choose between the status quo and an alternative. The first DCE study concerning insulin therapy is Aristides et al. (2004) who compared Humalog Mix 25, an insulin analogue, with rapid-acting human insulin Humulin 30/70 and found significant WTP in five European countries. Hauber et al. (2009) elicited preferences in a DCE for oral diabetes treatment in type 2 patients through a web-enabled survey. Special emphasis was on causes for non-adherence. Guimarães et al. (2009a) and Guimarães et al. (2009b) investigated preferences for oral versus injectable insulin therapy in a DCE. They found that once the psychological barrier to initiating insulin therapy had been overcome, patients accommodated and accepted injectable therapy as a treatment option.

2.3 The Experiment

Sample and Interview Strategy

This DCE was conducted in Germany in the Fall of 2007. Because one of the research questions is whether financing insulin analogue through contributions to statutory health insurance GKV or through copayment makes a difference in terms of preferences, only adult GKV members (some 90 percent of the population) were asked to participate. A professional market research institute specialized in health care issues was commissioned to recruit individuals and to perform the interviews, which were face-to-face by trained field investigators. Interviewers found participants mainly through their private contacts with people regularly taking part in surveys. Out of the total 1,110 respondents, 602 do not suffer from diabetes, 202 suffer from type 1, and 306 from type 2 diabetes. Within the type 2 diabetics group, a distinction is made between insulin-naive and insulin-treated patients (152 and 154 respondents, respectively). Diabetics are oversampled to be able to study heterogeneity in preferences. While the sample design allocated the non-diabetics randomly across the 16 Länder (states), ages, and gender, it distributed the type 2 diabetics equally

over the three age groups, 46-55, 56-65, and over 65 because type 2 diabetes occurs almost exclusively past age 45. The minimum duration of diabetes treatment (insulin injections or oral therapy) was six months. Because it is very difficult to find patients suffering from type 1 diabetes, randomization was limited to the 16 Länder in this case.

Questionnaire

The questionnaire is divided into four parts. Part 1: The interview begins with questions about the respondent's health (general health status, regular consumption of pharmaceuticals, chronic illness, body mass index (BMI), diabetes) and health insurance (such as yearly contribution or supplementary insurance). This part is the same for all participants. Part 2: This part of the survey distinguishes between non-diabetics, insulin-treated diabetics, and insulin-naive diabetics. For non-diabetics it contains detailed information about diabetes and its treatment. Respondents are asked to indicate their (subjective) probability of becoming insulin-dependent during their lifetime (using a visual analog scale). Patients treated with insulin are asked about the course of their disease, their insulin treatment, and its side effects. Insulin-naive patients are presented with information about diabetes and its treatment as well. They are asked how long they have suffered from diabetes, their treatment, and side effects. They are made to indicate their (subjective) lifetime probability of depending on insulin (again using a visual analog scale). Part 3: This part is the same for all participants. To prepare them for the DCE, the attributes are explained in detail, with special emphasis on the two payment vehicles "copayment" and "increase in contribution to health insurance". Since the interviews were face-to-face, respondents had the possibility to ask questions, and interviewers to offer more explanation. Then, the insulin used in current therapy is described to respondents (status quo). Eight times, an alternative type of insulin with changed attribute levels (alternative, see below) was presented and respondents asked to choose between the alternative and the status quo. Part 4: The interview finishes with socioeconomic items (gender, age, education, and residence). The last question is monthly household income to be

indicated on a visual analog scale to ensure a high response rate.

Attributes

Although both rapid- and long-acting insulin is required for successful therapy, this study only considers long-acting insulin. Current treatment guidelines use long-acting NPH insulin to provide base-level supply. This therapy constitutes the fixed status quo. It is defined by four attributes, which serve to reflect the differences in the properties of long-acting NPH insulin and insulin analogue. In contradistinction to other DCEs, no pretest was therefore necessary to establish the relevant attributes. They are the following.

Risk of hypoglycemia (*Hypo*, see Table 2.1) is one of the main side effects of insulin therapy. Its incidence depends on the individual, the dose of insulin needed, individual habits, and the insulin preparation. On average the number of hypoglycemic events can be estimated at 1 to 2 per week (Sreenan et al., 2008 and discussions with diabetologists). With a time horizon of up to six months (see weight attribute below), this puts the risk at 100 percent in the status quo. Most studies suggest that incidence is lower with insulin analogue than with NPH insulin (see Vague et al., 2002, Hermansen et al., 2004, Home et al., 2004, Kolendorf et al., 2004, Robertson et al., 2004, Russell-Jones et al., 2004, Russell-Jones, 2007, Dornhorst et al., 2008, Marre et al., 2009 and for meta-analyses Raskin, 2007, Satish and Ramachandra, 2008, Demssie et al., 2009, Freeman, 2009, Hermansen et al., 2009, and Monami et al., 2009). A study that does not find any differences in the frequency of hypoglycemia compared to NPH insulin is Umpierrez et al. (2009), while Singh et al. (2009) report mixed results. A Cochrane review (Horvath et al., 2007) concluded fewer analogue users experienced symptomatic overall or nocturnal hypoglycemic episodes compared to NPH insulin users. The magnitude of the decrease varies across studies. Hermansen et al. (2009) found a reduction of total hypoglycemic events of over 50 percent, Kolendorf et al. (2004) of 18 percent, and Vague et al. (2002) of 22 percent. IQWiG wrote in its final report (IQWiG, 2009) that insulin analogue

Table 2.1: DCE attributes, labels and levels

Overall risk of hypoglycemia (*Hypo*)	Status quo:	100%
	Alternative:	100% / 75% / 50%
Weight change (*Weight*)	Status quo:	+ 2,5 kg
	Alternative:	+ 2.5 kg / ± 0 kg / – 1.0 kg
Swinging before injection (*Swing*)	Status quo:	Necessary
	Alternative:	Necessary / Not necessary
Time of injection (*Flexibility*)	Status quo:	Predetermined
	Alternative:	Predetermined / Not predetermined
Copayment per year (*Copayment*)	Status quo:	None
	Alternatives:	None / Euro 50 / Euro 150 / Euro 300
Health insurance contribution (*Contribution*)	Status quo:	Individual contribution
	Alternatives:	± 0% / +0.5% / +1.0% / +2.0%

Note: Euro 1 ≈ USD 1.4 at 2008 exchange rates

significantly lowers the risk of severe (analogue: 0.0 percent vs. NPH: 2.1 percent), of mild (analogue: 57.0 percent vs. NPH: 78.2 percent, OR = 0.37), and of nocturnal hypoglycemia (analogue: 26.2 percent vs. NPH: 44.1 percent, OR = 0.45) for type 2 diabetes. A conservative value of 30 percent risk reduction is therefore attributed to insulin analogue. In order to have sufficient spread for statistical inference, the alternative incidence levels are set to 75 and 50 percent relative to NPH insulin in the DCE.

Obesity (*Weight*) is a major problem of type 2 diabetes patients. 80 percent suffer from obesity according to Russell-Jones and Khan (2007). Correspondingly, Häussler et al. (2005) found a significantly higher BMI in type 2 patients than in the overall German population. Insulin therapy makes this problem even worse. As a side effect of treatment with human insulin, patients gain weight, especially during the first months of insulin therapy. The UK Prospective Diabetes Study (UKPDS) Group (1998) observed a 2.5 kg increase over 6 months on average; this value serves to describe the status quo. Insulin analogue is found to mitigate weight gain (see Haak et al., 2003, Haak et al., 2005, Hermansen and Davies, 2007, Raslová et al., 2007, Russell-Jones

and Khan, 2007, Dornhorst et al., 2008, Demssie et al., 2009, Freeman, 2009, Mandosi et al., 2009, Marre et al., 2009, Monami et al., 2009). It may even cause weight loss of up to 1 kg (Russell-Jones, 2007, Sreenan et al., 2008, Hermansen et al., 2009, for meta-analyses see Bush, 2007, Raskin, 2007, and Satish and Ramachandra, 2008). The evidence allows to associate insulin analogue with a weight gain of 0 kg, while the levels used in the DCE are + 2.5, 0, and - 1 kg, respectively.

Before every injection, human NPH insulin has to be swung (not shaken) to achieve uniform dilution (*Swing*, see Table 2.1), ensuring injection of an optimal amount of insulin. This defines the status quo. Insufficient swinging causes a risk of injecting a suboptimal amount of insulin and inadequate control of blood sugar levels (Schleser-Mohr, 2007). Insulin analogue can be injected immediately, without swinging (Schmeisl, 2009). These two levels also appear in the DCE.

Another difference in the two types of insulin is flexibility with regard to time of injection (*Flexibility*, see Table 2.1 again). Human insulin reaches its maximum effect often after a few hours (Soran and Younis, 2006). The time of the bedtime injection therefore is set at 10 pm to avoid insufficient insulin levels in the early morning; this defines the status quo. Insulin analogue has a different action profile. Its maximum effect occurs later (see Kurtzhals, 2007 and Demssie et al., 2009), allowing patients to inject insulin already before 10 pm, usually between dinner and bedtime. However, time of injection should not vary from day to day. Insulin analogue is therefore described accordingly, and this attribute again has two levels in the DCE.

The last two attributes listed in Table 2.1 describe two modes of financing, individually through copayment by diabetics themselves or collectively through increased GKV contributions by the whole population. Inclusion of these two price attributes can be justified for at least three reasons. First, Germany has been introducing copayment on pharmaceuticals along with reference pricing of drugs, making it a mode of financing of increasing importance. Second, a population may well have preferences with regard

to modes of financing, as evidenced by Skjoldborg and Gyrd-Hansen (2003) for the case of Denmark. And third, economic considerations lead one to suspect that those affected by the disease prefer financing through increased insurance contributions (which fall on everyone) over copayment (which burdens only the affected). This hypothesis will be tested (see H4 of Section 2.5). As to *Copayment*, there is none in the status quo for diabetes patients, regardless of type of therapy (see Table 2.1). In the alternative, the levels are Euro 50, 150, and 300 per year, respectively. As to *Contribution*, respondents were asked to look up the actual amount paid to establish an individual-specific status quo. Contributions are estimated to increase by Euro 8.54 per year and GKV member[3] if insulin analogue is added to the benefit list. On average this corresponds to an increase of 0.5 percent of annual health insurance contributions, which is the value attributed to insulin analogue. In the DCE, levels characterizing the alternative are set to increases of 0.5, 1, and 2 percent, respectively.

Pretest and Design

The pretest was conducted by the same market research institute and consisted of 30 face-to-face interviews with individuals from the greater Leipzig area (17 non-diabetics, four type 2 insulin-dependent, four type 2 insulin-naive, and five type 1 patients, 23 women and 7 men, 52 years of age on average). One-third of the interviews were monitored by the first author of this study. In general, participants and interviewers understood the questions well. 25 individuals rated the choices "easy" and five "difficult". However, no one rated them "very difficult". On average the new insulin was chosen 3.8 times out of 10 choices. Econometric estimates confirmed the relevance of attributes and levels, with one exception. In the pretest, increases in insurance contributions were 0.25, 0.5, and 1 percent. Apparently, this range was not sufficient to affect decisions. Therefore it was scaled up to 0.5, 1, and 2 percent. Figure 2.1 shows an example of a choice question.

[3]On average, extra cost of treatment with insulin analogue rather than human insulin is Euro 226 per year and diabetic. Multiplied by the number of insulin-treated diabetics in Germany (=1.9 mn., see Giani et al., 2004) and divided by the number of GKV members paying contributions (=50.471 mn., see Federal Ministry of Health, 2008) one obtains Euro 8.54 per year and GKV member.

Choice Question: Would you prefer insulin-dependent diabetics to be treated with the current or the new insulin?

	Current Insulin	New Insulin
1 Events of hypoglycemia	on average 1-2 per week	approx. 25% lower risk
2 Weight change during first 6 months of therpay	+ 2.5kg weight gain	+ 2.5kg weight gain
3 Accuracy of dosage / preparation of insulin before every injection	Before every injection swinging necessary	No swinging necessary
4 Point in time of injection	Predetermined: After 10pm (daily identical)	Predetermined: After 10pm (daily identical)
5 Additional copayment per year	None	50 Euro
6 Your contribution to statutory health insurance per year	_____ Euro	+ 0.5% = +_____ Euro
In this situation I choose	☐ the current insulin	☐ the new insulin

Figure 2.1: Choice question example: Fixed status quo (current insulin) vs. alternative (new insulin)

For the main survey, a D-optimal design was constructed (Atkinson and Donev, 1992, Street et al., 2001, Burgess and Street, 2003, and Carlsson and Martinsson, 2003), using the software GOSSET (see Kuhfeld et al., 1994 and Sloane and Hardin, 2007). Out of the 576 possible combinations, 30 were retained in this way and divided into four card sets. Each set consisted of eight choices between the current insulin (status quo) and a new insulin (alternative). Consistency was tested by including weakly dominated alternatives, which however were favored only by a few respondents. "Expensive" alternatives were chosen significantly less often than "cheaper" ones. In total, the new insulin was picked in 40 percent, the current insulin in 60 percent of cases. 27 individuals did not alternate between the current and new insulins. Half of the respondents stated that decisions were "easy", 39 percent "difficult", and 11 percent "very difficult".

2.4 Ex-ante vs. Ex-post Willingness-To-Pay

Based on random utility theory (see Luce, 1959, McFadden, 1974, Manski and Lerman, 1977, McFadden, 1981, and McFadden, 2001), DCEs are designed to investigate individuals' preferences for (non-)marketed goods or goods that do not yet exist (Becker

and Zweifel, 2008). In a DCE, participants are repeatedly asked to choose between a fixed status quo and an alternative whose attributes take on different values each time. When choosing between alternatives, a rational individual will always select the alternative with the higher level of expected utility. Thus, neglecting the expectation operator for simplicity, the decision-making process functions as a comparison of utility values determined by

$$U_{ij} \equiv v(a_j, p_j, y_i, s_i, \varepsilon_{ij}), \tag{2.1}$$

where U_{ij} represents the indirect utility value attained by individual i in alternative j. It depends on the vector of attributes a_j, the price p_j, the individual's income y_i, and socioeconomic characteristics denoted by s_i. Finally, ε_{ij} is an error term that varies over alternatives and individuals. Provided the error term is additive, the individual will choose alternative k over alternative l if

$$u(a_k, p_k, y_i, s_i) + \varepsilon_{ik} \geq u(a_l, p_l, y_i, s_i) + \varepsilon_{il}, \tag{2.2}$$

where $u(\cdot)$ is the deterministic and ε_{ij} the stochastic component of the utility function $v(\cdot)$. The probability of choosing the alternative k over l, P_{ik}, is assumed to equal the probability of the difference in Equation (2.2) occurring. Solving for the difference in error terms, one obtains

$$P_{ik} = Prob[\varepsilon_{il} - \varepsilon_{ik} \leq u(a_k, p_k, y_i, s_i) - u(a_l, p_l, y_i, s_i)]. \tag{2.3}$$

For any inference about the left-hand side of inequality (2.3), a probability law for $\omega{=}(\varepsilon_{il} - \varepsilon_{ik})$ must be assumed. Since the logistic distribution assumes independence of irrelevant alternatives (IIA), the normal distribution is used here, resulting in probit estimation. It is assumed that errors are correlated between the choices of a given respondent but not across respondents, calling for random effects specification. With the utility function linear in parameters (Louviere et al., 2000), one has

$$\Delta U_{ik} = \beta_0 + \beta_1 a_{1k} + \beta_2 a_{2k} + \ldots + \beta_L a_{Lk} + \omega_{ik}, \tag{2.4}$$

with ω_{ik}=μ_i+ν_{ik}. Here, a_{1k}, ..., a_{Lk} are the L attributes of the alternative k in consideration. According to Equation (2.3), only differences in utility matter. For this reason, fixed characteristics of respondents drop out. The βs are the parameters to be estimated.

Based on Hanemann (1983), the marginal rate of substitution (MRS) between the two attributes m and n is equal to the ratio of the derivatives of the indirect utility function with respect to the two attributes,

$$MRS = \frac{\partial v/\partial a_m}{\partial v/\partial a_n} = \frac{\beta_m}{\beta_n}. \tag{2.5}$$

Defining n as a financial attribute allows interpretation of the negative of the MRS as a marginal WTP for attribute m.

A special feature of this study is that it seeks to measure WTP of both individuals who do not suffer from the disease or do not need insulin yet (ex-ante) and insulin-treated diabetes patients (ex-post). Whereas the utility gained (or lost) from a change in treatment is a real and immediate utility change for insulin-treated diabetics, it is an expected utility for non- and insulin-naive diabetics, which can be written as

$$EU_{ij} = \pi_i \cdot U_{ij}(\text{Therapy}|\text{Diabetic}) + (1 - \pi_i) \cdot U_{ij}(\text{Therapy}|\text{Non-Diabetic}), \tag{2.6}$$

where π_i is the individual-specific (subjective) probability to come down with insulin-treated diabetes. For patients treated with insulin, π_i is equal to 1, causing the second term of Equation (2.6) to become zero. In this case, Equation (2.6) is equal to U_{ij}, the individual's utility experienced form alternative j. When substituting the attributes described above into Equation (2.1), and assuming linearity, utility for insulin-dependent diabetics becomes

$$\begin{aligned} U_{ij} \quad = \quad & \beta_0 + \beta_1 Hypo_{ij} + \beta_2 Weight_{ij} + \beta_3 Swing_{ij} + \beta_4 Flexibility_{ij} \\ & + \beta_5 Copayment_{ij} + \beta_6 Contribution_{ij} + \omega_{ij}. \end{aligned} \tag{2.7}$$

For individuals not suffering from the disease and insulin-naive diabetics, π_i is between zero and one. Their expected utility function therefore reads,

$$\begin{aligned} EU_{ij} &= \pi_i \cdot (\beta_0 + \beta_1 Hypo_{ij} + \beta_2 Weight_{ij} + \beta_3 Swing_{ij} \\ &\quad + \beta_4 Flexibility_{ij} + \beta_5 Copayment_{ij} + \beta_6 Contribution_{ij}) \\ &\quad + (1 - \pi_i) \cdot (\beta_0 + \beta_6 Contribution_{ij}) + \omega_{ij}. \end{aligned} \tag{2.8}$$

Recall that the variables in Equation (2.7) represent the differences between the current and the new insulin. For example $Hypo_{ij}$ is the probability of suffering from hypoglycemia when treated with the current insulin minus this probability when treated with the new insulin (NPH insulin). Consequently, the values for $Hypo$, $Weight$, $Swing$, $Flexibility$, and $Copayment$ are set equal to zero in case of non-diabetics and insulin-naive patients because they do not vary across alternatives. However, health insurance contributions do vary since if the pharmaceutical is paid for by the GKV, every member contributes to the cost of the medications covered, not only patients.

There are two main reasons for a non-diabetic person to derive utility from and hence have a positive WTP for diabetes treatment, namely altruism and/or buying a call option for better treatment in case of coming down with the disease. Starting with the latter, the first term of Equation (2.8) shows the change in expected utility of a person who envisages coming down with insulin-dependent diabetes and therefore has positive WTP for a call option on new treatments. The higher the probability π_i, the higher the probability of exercising this option, and the higher WTP. With regard to altruism, the second term of Equation (2.8) represents the change in expected utility of a person who envisages staying healthy. In this case, β_0 can be interpreted as WTP due to altruism. Finally, Equation (2.8) can be rewritten as

$$\begin{aligned} EU_{ij} &= \beta_0 + \pi_i \beta_1 Hypo_{ij} + \pi_i \beta_2 Weight_{ij} + \pi_i \beta_3 Swing_{ij} \\ &\quad + \pi_i \beta_4 Flexibility_{ij} + \pi_i \beta_5 Copayment_{ij} + \beta_6 Contribution_{ij} + \omega_{ij}. \end{aligned} \tag{2.9}$$

This equation holds for non-diabetics as well as for diabetics. For the latter, π_i equals 1 if treated with insulin, causing Equations (2.9) and (2.7) to be identical. The calculation of WTP has to be modified as well. If the financial attribute (n) is specified to be copayment, Equation (2.5) holds. However, if it is GKV contributions, the probability of becoming a diabetic has to be taken into account,

$$\text{WTP} = -\pi_i \cdot \frac{\beta_m}{\beta_6}. \tag{2.10}$$

2.5 Hypotheses

This section is devoted to the statement of hypotheses concerning WTP values.

Hypothesis H1: From the GKV members' point of view, insulin analogue generates an additional utility compared to human insulin.

Increases in contributions and copayment will always have a negative effect on utility. However, this hypothesis states that the other attributes generate enough additional utility compared to human insulin to make its total effect positive.

Hypothesis H2: WTP values for the attributes are in the following rank order.

H2.1 Decreasing the risk of hypoglycemia has the highest WTP, followed by avoiding weight gain.

H2.2 WTP for more flexibility with regard to time of injection is considerably lower than for avoiding weight gain.

H2.3 WTP for no need to swing the preparation before injection is very low, not significantly different from zero.

Hypoglycemia is a traumatic experience. Symptoms of hypoglycemia include shakiness, dizziness, confusion, and difficulty to speak, just to mention a few. Severe

hypoglycemia can cause loss of consciousness and even death. Therefore the highest WTP is expected for a decrease in this risk, dominating concerns about weight gain. This is supported by Hermansen and Davies (2007), who found that patients often take a precautionary snack to avoid hypoglycemia, accepting weight gain as the consequence. Further supporting references are Guimarães et al. (2009b) (in the context of oral and inhaled insulin delivery) and Hauber et al. (2009) (in the context of oral glucose-lowering medications) who conclude that patients of both type 1 and type 2 have a higher WTP for avoiding hypoglycemia than for avoiding weight gain. In turn, avoiding weight gain is expected to generate a higher WTP than more flexibility with regard to time of injection. Aristides et al. (2004) analyzed WTP for flexibility in meal-time insulin injections. Whereas WTP values are significantly positive, they are lower than for avoiding weight gain as estimated by Guimarães et al. (2009b) and Hauber et al. (2009). Finally, failure to swing the preparation might be a worry for patients at the beginning of the treatment. With increasing experience permitting them to save time and effort, WTP for this attribute is predicted to go to zero. Recall that diabetics participating in the DCE had been subject to the condition for six months or more.

Hypothesis H3: There is significant heterogeneity of WTP values between diabetics and non-diabetics and between diabetes subgroups.

The difference in experience with using insulin might be the key reason for heterogeneity in preferences (as found in Guimarães et al., 2009b). Whereas type 1 and insulin-treated type 2 diabetics have used insulin before, non-diabetics and insulin-naive type 2 diabetics have not. For instance, they do not know what a hypoglycemic situation feels like and how it can be handled.

Hypothesis H4: Non-affected respondents and diabetics not treated with insulin prefer financing through patients themselves in the guise of copayment, whereas insulin-treated patients prefer financing

THROUGH HEALTH INSURANCE CONTRIBUTIONS.

Both diabetics and non-diabetics are predicted to have positive WTP for insulin analogue. However, WTP values of non-diabetics and insulin-naive diabetics are expected to be higher when financing occurs through copayment by patients themselves than jointly by the whole population through health insurance contributions. Conversely, WTP values of type 1 and insulin-dependent type 2 diabetics should be higher when financing occurs jointly through health insurance contributions.

2.6 Data: Descriptive Statistics

Table 2.2 gives an overview of the sample. Approximately 50 percent of the respondents are female. Average age is higher for type 2 diabetics than for the rest of the sample because this disease occurs primarily among the elderly (although the number of children suffering from type 2 diabetes has been increasing substantially, see Wabitsch et al., 2004 and Rosenbauer and Stahl, 2010). Respondents were asked to mark their subjective health status on a visual analog scale ranging from 0 (very bad health) to 100 (very good health). Non-diabetics reported the highest average value of 73, insulin-treated type 2 patients the lowest of 54. On average, type 2 diabetics have the highest BMI with 28 (insulin-treated) and 27 (insulin-naive), respectively. This matches the findings of the UK Prospective Diabetes Study (UKPDS) Group (1998) stating that obesity is highly prevalent among type 2 diabetics. The difference in BMI between type 2 and non-diabetics is statistically significant.

Average net household income is Euro 1,904 per month. Insulin-naive diabetics of type 2 have a lower income (Euro 1,783) than non-diabetics (Euro 1,975). This difference is in accordance with Häussler et al. (2005) who found a negative correlation between prevalence of type 2 diabetes and income. Because contributions to statutory health insurance GKV are defined as a percentage of (labor) income, higher incomes lead to

higher contributions. While the function is nonlinear because the percentage varies between sick funds and regions, non-diabetics do pay higher contributions on average than the others. Some 41 percent of them also have at least one supplementary insurance contract, compared to 30 percent for type 1 diabetics and 31 percent for insulin-treated diabetics. This reflects the fact that diabetics treated with insulin present high risks to private health insurers offering supplementary coverage, causing high premiums or exclusion clauses to be applied.

The lower part of Table 2.2 contains information about duration of illness and incidence of diabetic complications. Type 1 diabetics on average have been suffering for 17 years from the disease at the time of the DCE. For type 2 diabetics this value drops to 8 to 9 years. Only 18 percent of type 2 diabetes patients with insulin treatment do not suffer from any complication. For insulin-naive type 2 diabetics, this number is 23 percent and for type 1 diabetics, 27 percent. High blood pressure is the most common complication, followed by diabetic neuropathy, diabetic feet, and diabetic retinopathy. Strokes, hearth attacks, as well as amputations, are most common among type 2 diabetics with insulin therapy.

Table 2.2: Descriptive statistics

Variable	All respondents		Non-diabetics		Type 1 diabetics		Type 2 diabetics insulin-treated		Type 2 diabetics insulin-naive	
N	1,110		602		202		154		152	
Socioeconomic variables and health status										
Age	51.10	(16.18)	47.70	(16.54)	44.67	(15.17)	61.99	(9.74)	62.11	(9.44)
Female*	51.49		52.25		50.49		50.65		50.65	
Subjective health status[1]	66.46	(23.27)	72.56	(22.46)	62.06	(22.63)	53.70	(21.95)	61.33	(20.74)
BMI[2]	26.17	(4.54)	25.26	(4.16)	26.45	(5.32)	28.18	(4.50)	27.35	(3.98)
Health insurance										
Income[3]	1,903.75	(1,014.85)	1,974.55	(1,055.40)	1,814.07	(1,022.41)	1,866.67	(918.88)	1,783.22	(918.07)
GKV contribution[4]	1,879.60	(703.91)	1,914.77	(727.90)	1,832.82	(719.00)	1,894.09	(630.44)	1,787.76	(650.30)
Supplementary insurance*	37.03		40.51		30.20		31.17		38.16	
Duration of illness and incidence of diabetes complications										
Years of illness					17.32	(14.40)	8.60	(5.78)	8.03	(8.21)
Diabetes complication*					72.87		81.82		76.97	
High blood pressure*					43.07		63.64		59.21	
Diabetic foot*					20.30		30.52		16.45	
Diabetic neuropathy*					32.67		35.06		27.63	
Diabetic retinopathy*					10.89		14.94		7.24	
Stroke / heart attack*					8.91		12.39		5.26	
Amputation*					1.49		3.90		1.32	

[1]: Subjective health status from 0 = "very bad" to 100 = "very good"

[2]: Body mass index, according to the World Health Organization (2010) a BMI over 25 signals overweight

[3]: Net income per household per year in Euro

[4]: Health insurance contribution per year in Euro

Note: *: In percent of the respective subsample; standard deviations in parentheses

2.7 Empirical Results

2.7.1 Willingness-To-Pay

As a first step, it is important to know whether the attributes retained are relevant and have the expected impacts on utility. Table 2.3 presents the estimation results of Equation (2.9). All coefficients are highly significant and have the expected signs. The positive value of the constant can be interpreted as follows. If the specification of the utility function had been perfect, then the difference between the alternative and the status quo would be entirely due to the differences in attributes. There would be no reason to expect a constant different from zero. However, there may be individual characteristics not accounted for that give rise to a bias in favor or against the status quo (Salkeld et al., 2000). In the present case, the positive constant points to a preference in favor of the alternative and hence a bias against the status quo.

Using Equations (2.5) and (2.10), marginal WTP values depending on the mode of financing (copayment and increase in contributions, respectively) can be estimated. The upper part of Table 2.4 shows the results for copayment, the lower, for contributions. According to Equation (2.10) WTP values for the latter must be probability-weighted

Table 2.3: Results of a random-effects probit estimation, aggregate sample

Attribute	Expected sign	Coefficient	z-value	Marginal effect
Constant		0.7632	15.77	
Hypoglycemia[1]	+	0.0065	14.07	0.002
Weight[2]	+	0.1380	13.27	0.051
Swing[3]	±	0.2947	8.41	0.108
Flexibility[3]	+	0.1704	4.94	0.063
Copayment	–	−0.0055	−39.97	−0.002
Contribution	–	−0.0047	−5.23	−0.002
σ_μ		0.51	19.93	
ρ		0.20	12.53	

[1]: Decrease of the risk of hypoglycemia
[2]: Avoiding weight gain
[3]: Dummy-variable, 0 = status quo, 1 = alternative

for deriving estimates that apply to GKV members in general, who would pay increased contributions. Estimates weighted by the average subjective probability of coming down with insulin-treated diabetes are displayed in the last two columns of Table 2.4. Subjective probabilities (π_i) were measured in the questionnaire using a visual analog scale from 0 percent (will never become insulin-treated diabetic) to 100 percent (will become insulin-treated diabetic with certainty). For diabetics already treated with insulin, π_i is equal to 1. The average value over all respondents ($\bar{\pi}$) is 53 percent.

For both modes, preference for the alternative is very high, viz. Euro 262 and Euro 162 per year. In most DCEs, status quo bias is negative, indicating resistance against change (see e.g. Telser and Zweifel, 2002 and Zweifel et al., 2006). In the case of diabetes treatment, respondents seem to be willing to pay for a shift away from the status quo.

As to the risk of hypoglycemia, respondents are willing to pay an estimated Euro 1.19 per year for a 1 percentage point reduction through copayment and Euro 1.39 through contributions. The second amount decreases to Euro 0.74 per year when weighted by average probability $\bar{\pi}$ (see lower part of Table 2.4). To avoid 1 kg of weight gain, respondents are willing to pay Euro 25 through copayment or Euro 16 through higher yearly contributions, respectively.

To compare the importance of the attributes, consider a 100 percent change. Although unrealistic in the case of hypoglycemia, it allows to compare WTP directly with the (0,1) attributes. For the risk of hypoglycemia, a 100 percent decrease has an approximate WTP of Euro 119 (copayment) and Euro 74 (contribution), respectively. For fully avoiding the average weight gain of 2.5 kg (see Section 2.3), which also amounts to a 100 percent change, the WTP value is Euro 63 ($= 2.5 \cdot 25.15$, copayment) and Euro 39 ($= 2.5 \cdot 15.55$, contribution). Hence, regardless of mode of financing, respondents value lowering the risk of hypoglycemia two times more than avoiding

weight gain, corroborating H2.1. As to WTP for increased flexibility with regard to the timing of the injection, the values amount to Euro 31 (copayment) and Euro 19 (contribution), respectively. This is much less than the Euro 63 and Euro 39 for avoiding weight gain, in accordance with H2.2.

The possibility to inject insulin without swinging before every injection is worth Euro 54 (copayment) or Euro 33 per year (contribution), respectively. Since these values differ from zero, they constitute evidence against H2.3. A seemingly minor innovation (from the medical point of view) is clearly valued by consumers. However, it is valued less than avoidance of either hypoglycemia or weight gain. For instance, the difference between Euro 119 (100 percent change in hypoglycemia, copayment) and Euro 54 (swing, copayment) has statistical significance in view of the small standard errors.

Table 2.4: Marginal WTP for product attributes, aggregate sample, Euro

Attribute	MWTP	Standard error Delta Method[4]	Bootstrap[5]	z-value	MWTP · $\bar{\pi}^*$
Financing through copayment					
Constant	261.50	8.54	9.11	30.62	
Hypoglycemia[1]	1.19	0.09	0.10	13.48	
Weight[2]	25.15	1.90	2.19	13.23	
Swing[3]	53.69	6.34	6.31	8.47	
Flexibility[3]	31.04	6.29	6.37	4.94	
Financing through health insurance contribution					
Constant	161.75	29.20	41.11	5.54	161.75
Hypoglycemia[1]	1.39	0.28	0.40	4.87	0.74
Weight[2]	29.25	5.79	8.80	5.05	15.55
Swing[3]	62.46	13.87	18.48	4.50	33.21
Flexibility[3]	36.11	10.20	13.31	3.54	19.20

* Except constant

[1]: Decrease of the risk of hypoglycemia by 1 percentage point

[2]: Avoiding weight gain

[3]: Dummy variable, 0 = status quo, 1 = alternative

[4]: Standard errors calculated using the delta method

[5]: Standard errors calculated using bootstrapping with 1,000 replications

Note: All marginal WTP (MWTP) values are in Euro per year, Euro 1 ≈ USD 1.4 at 2008 exchange rates

To test H1 (positive value of the new pharmaceutical) total WTP values need to be calculated. As described in Section 2.3, insulin analogue corresponds to the following changes in attributes. Risk of hypoglycemia decreases by 30 percent in comparison to treatment with human insulin NPH. Whereas patients gain 2.5 kg on average with human insulin, there is no weight change with insulin analogue. The preparation does not need to be swung, and the timing of injection is more flexible. Following Hanemann (1983), WTP associated with these non-marginal changes is computed as the marginal WTP multiplied by the change of the attribute's value. These component values are then summed up to obtain total WTP for the product (see Johnson and Desvousges, 1997). The results of these calculations are shown in Table 2.5. Total WTP for the new drug amounts to Euro 445 per year if financed through copayment and Euro 275 (probability-weighted) if financed through an increase in contributions. Approximately 60 percent of this WTP comes from bias in favor of the alternative. Even if this component is subtracted, the resulting values of Euro 183 and Euro 114, respectively, are still significantly positive in view of the small estimated standard errors displayed in Table 2.5. Therefore, H1 is confirmed.

Table 2.5: WTP for product attributes, aggregate sample, Euro

	Financing through copayment		Financing through contribution	
Attribute	WTP	z-value	WTP $\cdot\ \bar{\pi}^*$	z-value
Constant	261.50	16.29	161.75	5.54
Hypoglycemia[1]	35.74	13.48	22.20	4.87
Weight[2]	62.87	13.23	38.88	5.05
Swing[3]	53.69	8.47	33.21	4.50
Flexibility[3]	31.04	4.94	19.20	3.54
Total	444.84		275.24	
Total net of constant	183.34		113.49	

* Except constant

[1]: Decrease of the risk of hypoglycemia by 30 percent

[2]: Avoiding a 2.5 kg weight gain

[3]: Dummy variable, 0 = status quo, 1 = alternative

Note: All WTP are in Euro per year, Euro 1 $\approx$ USD 1.4 at 2008 exchange rates

2.7.2 Willingness-To-Pay across Subgroups

To obtain group-specific WTP values, Equation (2.9) is estimated separately for non-diabetics, type 1 diabetics, type 2 insulin-naive as well as for insulin-treated diabetics. Group-specific MWTP values (not shown) are multiplied by the changes in attribute levels due to insulin analogue and summed, in full analogy to Table 2.5. The subjective probability of acquiring insulin-treated diabetes is 26 percent on average for non-diabetics and 56 percent for insulin-naive patients. The resulting non-marginal WTP values across subgroups are presented in Table 2.6. Sum I comprises all component WTP values, Sum II only the significant ones. Standard errors (z-values shown) are small enough to conclude that there is preference heterogeneity between these four groups, confirming H3.

Moreover, comparison of the upper and the lower part of Table 2.6 shows that the mode of payment matters, but not entirely in the way predicted by H4. As stated by H4, WTP values among diabetics should be higher when the new pharmaceutical is financed through increased GKV contributions rather than copayment, while among the non-affected, it should be the other way round. Now non-diabetics indeed exhibit a higher total WTP value when financing is through copayment. They are joined by the insulin-naive diabetics who apparently deem themselves not to be affected. On the other hand, type 1 diabetics do have higher WTP when financing occurs through increased contributions, but the difference is not statistically significant. For insulin-treated type 2 diabetics, the ordering is as expected at first sight (Sum I). Their WTP is extremely high when they envisage financing through increased contributions rather than copayment. However, not a single component value is significantly different from

Table 2.6: WTP for product attributes, stratified by diabetes type

	Non-Diabetics		Type 1 diabetics		Type 2 diabetics insulin-treated		Type 2 diabetics insulin-naive	
Attribute	WTP	z-value	WTP	z-value	WTP	z-value	WTP	z-value
Financing through copayment								
Hypoglycemia[1]	38.53***	10.76	27.95***	4.56	29.25***	4.02	43.98***	5.69
Weight[2]	71.80***	11.23	37.49***	3.39	71.53***	5.35	50.16***	3.69
Swing[3]	56.62***	6.67	48.28***	3.28	72.17***	4.03	25.85	1.43
Flexibility[3]	25.22***	3.00	24.37*	1.65	50.71***	2.89	46.45***	2.57
Constant	597.47***	13.58	106.90***	5.50	94.62***	4.08	286.55***	6.42
Sum I	789.63		244.99		318.29		452.99	
Sum II[4]	789.63		244.99		318.29		427.14	
Financing through health insurance contributions								
Hypoglycemia[1]	11.32***	3.88	34.11*	1.89	100.11	0.63	17.65***	2.51
Weight[2]	21.09***	4.00	45.75*	1.88	244.76	0.65	20.13***	2.38
Swing[3]	16.63***	3.55	58.92*	1.80	246.97	0.64	10.38	1.28
Flexibility[3]	7.41**	2.41	29.74	1.28	173.53	0.62	18.64*	1.87
Constant	175.51***	4.38	130.46**	2.20	323.79	0.67	115.00***	2.93
Sum I	231.96		298.99		1,089.16		181.80	
Sum II[4]	231.96		269.25		0.00		171.43	

[1]: Decrease of the risk of hypoglycemia by 30 percentage points

[2]: Avoiding a 2.5 kg weight gain

[3]: Dummy-variable, 0 = status quo, 1 = alternative

[4]: Only significant values

Note: * indicates significance at the 10 percent level, ** at the 5 percent level, and *** at the 1 percent level, all WTP values are in Euro per year

zero, causing Sum II to be zero as well. Apparently, opinions concerning insulin analogue are very divided among these patients as soon as it were to be paid for by increased contributions.

The high WTP values estimated for non-diabetics in the case of copayment also merit discussion. It is doubtful that they would be verified in a real purchase decision. Rather, being importantly due to a high constant, they point to a strong bias in favor of the alternative - provided those affected pay for the new drug themselves.

Finally, the entries of Table 2.6 can also be interpreted in the following way. The high copayment-related WTP values of non-diabetics and insulin-naive diabetics suggest that they prefer financing through patients themselves. Conversely, insulin-treated patients prefer financing jointly through health insurance contributions. However, whatever the group considered and regardless of mode of payment, WTP for insulin analogue measured by Sum I exceeds its cost of treatment (estimated at Euro 226 per year). If measured by Sum II, this is also true, with the only exception of type 2 insulin-treated patients whose preferences are too heterogeneous. Therefore, by a benefit-cost criterion, including this product in the GKV list of benefits appears to be justified.

2.8 Conclusions

This study revolves around the issue of whether a particular new pharmaceutical should be included in the benefit list of a social health insurer. From a cost-benefit perspective and neglecting distributional concerns, inclusion is justified if the insured have a willingness-to-pay (WTP) that exceeds the cost of treatment with the new product. The case in question is modern insulin therapy, using the long-acting insulin analogue "Insulin detemir". Preferences for this preparation in comparison to conventional therapy (using human insulin) are derived with the help of a discrete-choice

experiment. It involved 1,110 members of German statutory health insurance (GKV) in 2007, of whom 202 suffer from type 1 diabetes, 154 from type 2 diabetes treated with insulin, 152 are insulin-naive type 2 diabetics, and 602 are non-diabetics. The novelty of the experiment lies in two aspects. First, distinguishing these groups allows to estimate both ex-ante WTP for non-diabetics and naive diabetics and ex-post WTP for diabetic patients treated with insulin. Second, including the mode of payment (copayment vs. increased GKV contribution) permits to test whether the new drug has a favorable benefit-cost ratio regardless of the way it is financed. Based on the results reported, four research questions can be answered.

(1) Is there positive WTP for the long-acting insulin analogue? The evidence suggests there is, compared to the conventional therapy using long-acting human insulin NPH (Table 2.5). Components of this total value are WTP for reduction of the risk of hypoglycemia by 30 percent, no weight gain rather than 2.5 kg during the first six months of the therapy, relief from the need to swing the preparation before each injection, and flexibility with regard to the timing of the injection.

(2) Which product attributes contribute to total WTP? All product attributes have positive estimated WTP values. For comparison purposes, a hypothetical 100 percent reduction of the risk of hypoglycemia and of the weight gain are considered because the other attributes are (0,1) variables. In accordance with expectations, the maximum WTP value comes from risk reduction with respect to hypoglycemia, followed by avoiding weight gain. The other attributes are less highly valued.

(3) Is there preference heterogeneity across morbidity groups, viz. non-diabetics, type 1 diabetics, insulin-treated type 2 diabetics, and insulin-naive type 2 diabetics? Estimates do point to heterogeneity. Total WTP values differ significantly between subgroups. Non-affected insulin-naive type 2 and non-diabetics have similar preferences, as do affected type 1 and insulin-treated type 2 diabetics.

(4) Is the benefit-cost ratio of the new pharmaceutical favorable regardless of whether it is financed jointly through increased GKV contributions or by patients themselves through copayment? The evidence suggests this to be the case, with the one exception of type 2 insulin-treated diabetics, whose WTP values are very high but lack statistical significance. Also, whereas non-diabetics and insulin-naive diabetics exhibit higher WTP values if financing is through copayment, insulin-treated diabetics have higher values if financing is through insurance contributions. This can be interpreted as a preference for financing through copayment on the part of the non-affected non-diabetics and insulin-naive diabetics and through insurance of the part of the affected insulin-treated diabetics. However, since even non-diabetics' WTP is higher than the actual treatment cost of insulin analogue regardless of mode of payment, its inclusion in the German statutory health insurance GKV list of benefits can be justified.

These conclusions are subject to a number of reservations. First, the WTP estimates may be biased upward because participants in the experiment may not be representative of the GKV population. Indeed, the average net household income in the sample is below average, which may result in a general dissatisfaction with the status quo. This might drive up WTP for alternative treatment of diabetes as well. Second, in spite of differentiating between disease-specific groups, there still may be hidden heterogeneity that could correlate with error terms, causing bias in estimates. Finally, one may judge the cost-benefit standard adopted here as inappropriate. On the one hand, benefits should be measured in terms of quality adjusted life years rather than money according to some writers (see e.g. Culyer, 1990, Williams and Cookson, 2000, or Drummond et al., 2005). On the other hand, average WTP values neglect distributional issues.

While these concerns may well be valid, they are unlikely to overthrow the major findings of this study. First, there is clear evidence suggesting that not only the avoidance of hypoglycemia and weight gain but also attributes that typically are judged medically irrelevant such as no need for preparation (swinging) and flexibility

with regard to the timing of the injection are valued attributes of insulin therapy. In addition, these attributes have positive WTP values among diabetes patients and potential patients alike. Second, these valuations add up to total amounts that exceed the marginal cost of the new drug, with the only exception of type 2 insulin-treated diabetics whose WTP estimates, while sizable, cannot be distinguished from zero due to excess heterogeneity. It is difficult to conceive of biases so strong and distributional weightings so skewed to conclude that WTP values of GKV members likely fail to justify inclusion of this new pharmaceutical in the benefit list.

Disclaimer

This study was paid for by Novo Nordisk Pharma GmbH. However, the authors independently designed the experiment, and analyzed and interpreted the results without any influence from the sponsor. The market research institute was selected and paid for by the authors and delivered the data directly to them.

Chapter III
Why the Linear Utility Function is a Risky Choice in Discrete-Choice Experiments

Michèle Sennhauser

Submitted to

Health Economics

3 Why the Linear Utility Function is a Risky Choice in Discrete-Choice Experiments

3.1 Introduction

Discrete-choice experiments (DCEs) enjoy great popularity with the number of applied studies increasing steadily and penetrating every branch of the health-economic field (see Louviere and Lancsar, 2009). Methods of DCEs are also improving. Improvements in designs, attribute choice, and questionnaire methods have all recently emerged. For example, Green and Gerard (2009) were the first to implement cost-effectiveness of alternatives in the attributes. At the same time, more complex estimation procedures are being used. Whereas logit or probit estimations are most commonly applied, random coefficient models (also called mixed logits) are becoming more frequently used (Regier et al., 2009).

One component of DCEs that remains unchanged during this process of improvement: The form of the utility function. Almost all studies refer to Louviere et al. (2000), who stated that a linear specification in linear models typically accounts for 70 to 90 percent of explained variance. Consequently, most authors choose a main effects design and assume that all interactions are equal to zero (Amaya-Amaya et al., 2008, for an example, see Skjoldborg and Gyrd-Hansen, 2003). Interactions are then implemented rather ad hoc, or in the guise of interactions with socioeconomic characteristics (see for example Gerard et al., 2008). However, it can be argued that the utility function is unlikely to be linear because of diminishing marginal utility and gain-loss asymmetries (Hoyos, 2010).

So far, there have only been small attempts in the DCE literature toward a non-linear specification. Ryan and Watson (2008) note that, at the design stage, the researcher should consider the form of the utility function, taking account of potential non-linearities. However, they then proceed only estimating main effects, with interactions and higher-order terms assumed as negligible with the justification to be consistent with most DCE applications. Lancsar and Louviere (2006) outline difficulties with assuming linear utility functions. Because tests for dominance and lexicographic preferences rely on this assumption to hold, respondents previously labeled as "irrational" may simply appear to be so due to the specification, but are in fact not. A number of studies go further by allowing for interactions. Linearity assumptions of particular attributes were tested using a Wald or a Likelihood-Ratio (LR) test (as for example Telser and Zweifel, 2002).

Among econometricians, there is an ongoing methodological debate about model specification, but so far there is no "best" way of finding a correct model. As Kennedy (2003) notes, however, the debate has given birth to the general principle that economic theory should be the foundation of the model. Simultaneously, the data should help create a "more informed" economic theory by using econometric misspecification tests. However, to the knowledge of the author, this way of specification has not been applied yet to DCEs.

This paper sheds some light on this issue by showing that the linear utility function can be a risky choice in DCEs. For this purpose, a DCE conducted by Becker (2006, Chapters 6-8) in Switzerland is reexamined. The experiment elicits willingness-to-pay (WTP) for debated options in Swiss mandatory social health insurance. The DCE is evaluated in two ways. First, the utility function is assumed to be linear in the attributes; second, a non-linear utility function is used, employing econometric misspecification tests (as outlined by Hosmer and Lemeshow, 2000). The results are compared in terms of goodness-of-fit and estimated WTP. The findings suggest that not only the non-linear function outperforms the linear specification with regard to

goodness-of-fit, but also generates significantly different WTP. The results conclude that the form of the utility function may have significant impact on estimated WTP. In order to produce unbiased estimates of preferences, the specification of the utility function should be given more attention in future experiments.

The remainder of this chapter is structured as follows. Section 3.2 outlines the theoretical foundations of DCEs, model specification, and assessing goodness-of-fit. Section 3.3 introduces the experiment, and outlines the attributes and the interview strategy. Section 3.4 presents the empirical results with regard to the specification of the utility function and comparison of goodness-of-fit and WTP. Section 3.5 concludes.

3.2 Theoretical Foundations

3.2.1 Discrete-Choice Experiments

Based on random utility theory (see Luce, 1959, McFadden, 1974, Manski and Lerman, 1977, McFadden, 1981, and McFadden, 2001), DCEs are designed to investigate individuals' preferences for (non-)marketed goods or goods that do not yet exist (Becker and Zweifel, 2008). In a DCE, participants are repeatedly asked to choose between a fixed status quo and an alternative whose attributes take on different values each time. When choosing between alternatives, a rational individual will always select the alternative with the higher level of expected utility. Thus, neglecting the expectation operator for simplicity, the decision-making process functions as a comparison of utility values determined by

$$U_{ij} \equiv v(a_j, p_j, y_i, s_i, \varepsilon_{ij}), \tag{3.1}$$

where U_{ij} represents the indirect utility value attained by individual i in alternative j. It depends on the vector of attributes a_j, the price p_j, the individual's income y_i, and socioeconomic characteristics denoted by s_i. Finally, ε_{ij} is an error term that varies

over alternatives and individuals. Provided the error term is additive, the individual will choose alternative k over alternative l if

$$u(a_k, p_k, y_i, s_i) + \varepsilon_{ik} \geq u(a_l, p_l, y_i, s_i) + \varepsilon_{il}, \tag{3.2}$$

where $u(\cdot)$ is the deterministic and ε_{ij} the stochastic component of the utility function $v(\cdot)$. The probability of choosing the alternative k over l, P_{ik}, is assumed to equal the probability of the difference in Equation (3.2) occurring. Solving for the difference in error terms, one obtains

$$P_{ik} = Prob[\varepsilon_{il} - \varepsilon_{ik} \leq u(a_k, p_k, y_i, s_i) - u(a_l, p_l, y_i, s_i)]. \tag{3.3}$$

For any inference about the left-hand side of Inequality (3.3), a probability law for $\omega{=}(\varepsilon_{il} - \varepsilon_{ik})$ must be assumed. Since the logistic distribution assumes independence of irrelevant alternatives (IIA), the normal distribution is used here, resulting in probit estimation. It is assumed that errors are correlated between the choices of a given respondent but not across respondents, calling for random effects specification. With the utility function linear in parameters (Louviere et al., 2000), one has

$$\Delta U_{ik} = \beta_0 + \beta_1 a_{1k} + \beta_2 a_{2k} + \ldots + \beta_L a_{Lk} + \omega_{ik}, \tag{3.4}$$

with $\omega_{ik}{=}\mu_i{+}\nu_{ik}$. Here, a_{1k}, ..., a_{Lk} are the L attributes of the alternative k in consideration. According to Equation (3.3), only differences in utility matter. For this reason, fixed characteristics of respondents drop out. The βs are the parameters to be estimated. With a non-linear utility function, interactions and higher-order terms of the attributes are also in Equation (3.4).

Based on Hanemann (1983), the marginal rate of substitution (MRS) between the two attributes m and n is equal to the ratio of the derivatives of the indirect utility function with respect to the two attributes. In the case of a linear utility function this is

$$MRS = \frac{\partial v / \partial a_m}{\partial v / \partial a_n} = \frac{\beta_m}{\beta_n}. \tag{3.5}$$

Defining n as a financial attribute allows interpretation of the negative of the MRS as a marginal WTP for attribute m. With a non-linear utility function, the MRS is no longer constant and can only be stated subject to the given values of the other attributes.

3.2.2 Specification of the Utility Function

As outlined in the introduction, economic theory should form the foundation of the utility function's specification. Indeed, misspecification tests from the econometrics and statistics literature should help to create the model. Hosmer and Lemeshow (2000, Chapter 4) provide an overview of these methods and present a strategy for binary response models. In the following, their 5-step procedure is summarized with regard to DCEs. Steps 1, 2, and 5 concern the issue of choosing variables that belong in the utility function. If too many variables are included, the problem of over-fitting arises, typically characterized by unrealistically large estimates of coefficients and/or standard errors (see Harrell et al., 1996). If an insufficient number of variables is included or variables that do not belong into the model are used, the generated predictions are also poor. Steps 3 and 4 concern the issue of choosing the attributes' functional form.

Step 1: As a first step, Hosmer and Lemeshow (2000) propose a careful univariate analysis of each of the possible covariates for the model. In the case of DCEs, these are the attributes. Contingency tables, smoothed scatter plots, and LR tests are some of the instruments that can be used. After the researcher determines a general impression

of the relations between the dependent variable (0 if the respondent decides in favor of the status quo, 1 if in favor of the alternative) and the independent variables (the attributes), a stepwise method may be applied to decide which attributes should be considered. It can either be a forward selection with a test of backward elimination or a backward elimination followed by a test for forward selection. According to Mickey and Greenland (1989), the significance level of entry into the model should not be equal to the traditional values (such as 0.05) because important variables could be excluded mistakenly. They recommend using a value between 0.15 and 0.25. However, most DCEs are designed and pretested in such a way that all attributes are important. Nevertheless, attributes can still find their way into the utility function despite insignificance if they are important for answering the research questions.

Step 2: The second step is to verify all the attributes that survived the selection procedure of Step 1. This should first include a Wald statistic for each variable. Attributes that do not contribute to the model are then excluded and the new model is compared to the former using an LR test. Estimated coefficients should also be compared when excluding an attribute. If they change markedly in magnitude, this indicates that the excluded variable was important for providing an adjustment of the effect of the attributes remaining in the model.

Step 3: Now that all attributes are verified, Hosmer and Lemeshow (2000) suggest exploring the scales of the continuous covariates. As a starting point, it is assumed that the utility function is linear in the attributes. There are different methods to ascertain this assumption, three of which follow. (1) A univariate smoothed scatter plot (Cleveland, 1979 and Cleveland and Devlin, 1988) shows potential non-linearities in the data and can easily be performed using statistical packages. (2) Four dummy variables ("design variables") are generated for the quartiles of the attribute. These are regressed together with the other attributes (but without the attribute in consideration and the first quartile's dummy) on the dependent variable. The quartiles' means are plotted against the estimated coefficients of the dummies. For the first quartile, the

coefficient is set to zero. The shape of this curve shows whether the linear specification might be appropriate. (3) The modified Hosmer-Lemeshow test (see Section 3.2.3 below), which can be applied for each variable separately. If this test fails, the linear specification is probably incorrect.

Step 4: In principle, the utility function needs to be as rich as the data requires, including the possibilities of interactions of higher order. However, all possible interactions should not be done in a model with many attributes, because this implies a very high-order regression. Consequently, the fourth step is to assess the need to include interaction terms. All possible interactions are tested using LR tests. To verify the results of Step 3, the terms in squares and other higher orders that resulted from Step 3 are tested as well.

Step 5: As a last step, in addition to assessing goodness-of-fit of the specified utility function, Hosmer and Lemeshow (2000) propose to backwardly select for more parsimony. However, the final backward selection has to carefully consider the fit of the model. If too many interactions and higher-order terms are dismissed, the utility function may no longer pass the goodness-of-fit tests.

3.2.3 Assessing Goodness-of-Fit

To decide which utility function is "better", linear or non-linear, a variety of goodness-of-fit measures are available. These include the LR test, the Akaike Information Criterion (AIC), the Bayesian Information Criterion (BIC), and the log-likelihood. However, the following tests for misspecification can also be used, as proposed by Basu et al. (2004) and Basu et al. (2006).

- Pregibon's Link test (see Pregibon, 1980 and Pregibon, 1981): This is a parsimonious test for non-linearity. Based on the initial estimate of the regression coefficients, a prediction of the dependent variable is generated. The prediction and the prediction squared are included as the only covariates in a second version

of the model. If the specification is truly linear, then the coefficient of the squared term should not be significantly different from zero.

- Ramsey's Reset test (see Ramsey, 1969): The Reset test allows for a richer form of model failure than the Link test. By including not only the prediction and the prediction squared but also the cube and prediction to the power of four, it allows for an s-shaped misfit compared to only a quadratic misfit with the Link test.

- Copas test (see Copas, 1983): This is a split-sample, cross-validation test for over-fitting the data. The sample is randomly divided into estimation data and test data. From a regression using the first sample, the predicted values are saved. These are used as the only covariate in a regression with the test data set. If the coefficient of the predictions is significantly different from one, over-fitting is a problem.

- Modified Hosmer-Lemeshow test (see Hosmer and Lemeshow, 2000): By observing the pattern in the residuals of the estimation as a function of the predicted values, this test determines whether there is a systematic bias. The modified Hosmer-Lemeshow test regresses the residuals on dummies for the deciles[1] of the predicted values. An F-test shows if the dummies have a significant influence on the residuals. If so, there is a non-linearity in the underlying data that is not represented in the model. The pattern of the regression coefficients and their standard errors allow a conclusion about the appropriate non-linear specification.

- Regular Hosmer-Lemeshow test (see Hosmer and Lemeshow, 1980): As with the modified version, the predicted values are grouped into ten equal sized groups. A Pearson-χ^2-statistic compares the observed and estimated expected frequencies and points out possible lacks of fit.

[1] Depending on the size of the data set, the test can be performed with more than ten groups.

- Pearson correlation: A Pearson correlation significantly different from zero between the residuals and the predicted values indicates that the model's predictions are biased.

3.3 The Experiment

3.3.1 Background: Swiss Statutory Social Health Insurance

The DCE assesses preferences for Swiss statutory health insurance and WTP for proposed reforms. It was conducted by Becker (2006, Chapters 6-8) in 2003.[2] Switzerland is a country of interest because its health insurance combines mandatory and choice elements in a way similar to the US and the Netherlands (OECD, 2004). The Health Insurance Law (KVG), effective since 1996, obliges all permanent residents of Switzerland to purchase health insurance policies for basic coverage. The law defines a uniform basic package of health care benefits that has become more comprehensive over the years, mostly driven by technological progress and new treatment methods. For a new therapy or pharmaceutical product to be included in the benefit package, its effectiveness, efficacy, and economic efficiency have to be proven (Article 32 Health Insurance Law KVG). Premiums are community-rated and not tax-financed, i.e. all insured pay approximately the same independent of age and morbidity. In addition, there is a fixed rate of copayment, amounting to 10 percent of health care expenditures with a maximum of CHF 600 per year (CHF 1 $\approx$ USD 0.82 at 2005 exchange rates).[3] Complementing these mandatory elements, there are elements of choice. There is free choice of health insurer. Contrary to the US, employers are not involved in this decision. Insurers are obliged by law to accept any applicant (for mandatory insurance, but not for supplementary health insurance). When it comes to choice of contract, there are two main elements. The first choice is the level of annual deductible. It ranges from a minimum of CHF 300 to a maximum of CHF 1,500.[4] The second involves

[2] See also Becker and Zweifel, 2008

[3] The maximum was increased to CHF 700 per year in 2004.

[4] This range was effective until 01.01.2005, when the maximum was increased to CHF 2,500 per year.

choosing between the conventional and Managed Care (MC) options. In the standard case, the choice of the provider is not restricted. In the MC settings, alternatives are offered. These include physician networks (similar to Independent Provider Associations in the US), restricted lists of physicians (Preferred Provider Organizations) and Health Maintenance Organizations (HMOs). For MC options, insurers are allowed to give reductions in premiums up to a certain percentage. However, the basic package of benefits remains the same, independent of the deductible and model chosen.

3.3.2 Attributes

The DCE's attributes represent different aspects of health insurance contracts within the context of Swiss health insurance (see Table 3.1). The importance of the attributes was secured by discussions with various experts from the Swiss health care system. Further data comes from a survey conducted by the Swiss Society of Applied Social Research (GFS, 2001). The first attribute of interest is reimbursement of alternative

Table 3.1: DCE attributes, labels and levels

Alternative medicine (*Alternative*)	Status quo:	Some alternative treatments are covered
	Alternative:	More alternative treatments are covered
Pharmaceuticals (*Generics*)	Status quo:	All pharmaceuticals on the list of benefits are reimbursed
	Alternative:	Only the cheapest product is reimbursed
Access to treatment methods (*Wait_Innovation*)	Status quo:	Coverage as soon as approved
	Alternative:	Coverage only two years after approval
Copayment per year (*Copayment*)	Status quo:	10 percent with a max. of CHF 600
	Alternative:	20 percent with a max. of CHF 1,200
Deductible per year (*Deductible*)	Status quo:	CHF 230, 400, 600, 1,200, 1,500
	Alternatives:	CHF 0, 2,400, 4,800
Health insurance contribution per month (*Contribution*)	Status quo:	Individual contribution
	Alternatives:	CHF –50, –25, –10, +10, +25, +50

Note: CHF 1 ≈ USD 0.82 at 2005 exchange rates

medicine (*Alternative*). In the status quo insurance contract, acupuncture, traditional Chinese medicine, anthroposophic medicine, homeopathy, neural therapy, and phytotherapy are part of the benefit package. The alternative suggests more treatments be reimbursed, such as treatments of alternative practitioners and naturopathy. The second attribute is reimbursement of pharmaceuticals (*Generics*). Whereas in the status quo insurance contract, where all pharmaceuticals on the list of benefits are reimbursed, the alternative offers only the cheapest product (the generics) to be paid by the health insurer. Another constraint is access to treatment methods (*Wait_Innovation*). While in the status quo contract, access is guaranteed to all insured immediately after approval, the alternative allows coverage only after 2 years. Two issues of great interest in the ongoing reform debate are consumers' willingness-to-accept copayment (*Copayment*) and deductibles (*Deductible*). With a copayment rate of 10 percent and a maximum payment of CHF 600 per year, the status quo's copayment rate is lower than the alternative's, which offers a 20 percent rate and a maximum of CHF 1,200 per year. For the deductible, there are five options, ranging from CHF 230 up to CHF 1,500 per year. The alternatives offer a wider range, starting from no deductible at all to CHF 4,800 per year. The sixth attribute is health insurance contributions (*Contribution*). The amount of contribution varies by increases and decreases of up to CHF 50 per month in the alternatives.

3.3.3 Pretest and Design

These and other attributes were checked for relevance in a pretest conducted with 20 individuals. The respondents understood the survey well and did not find the DCE section difficult. However, an adjustment had to be made with the deductible. While the levels for the alternative were CHF 3,000 and CHF 6,000 in the pretest, these were lowered to CHF 2,400 and CHF 4,800 to avoid protest response for lack of realism. For the main survey, the number of possible scenarios was reduced from a full factorial design to a fractional factorial D-optimal design (see Atkinson and Donev, 1992, Street et al., 2001, Burgess and Street, 2003, and Carlsson and Martinsson, 2003) of 27 choice sets using the program GOSSET (see Kuhfeld et al., 1994 and Sloane and Hardin,

Choice Question: Which insurance contract do you prefer?

	Current Contract	New Contract
1 Reimbursement alternative medicine	Some methods are covered	More methods are covered
2 Reimbursement pharmaceuticals	All pharmaceuticals are covered	Only the cheapest product
3 Access to new treatment methods	Coverage as soon as aprooved	Coverage as soon as aprooved
4 Copayment per year	10 % with a max. of 600 CHF	20 % with a max. of 1,200 CHF
5 Your deductible per year	_____ CHF	1,500 CHF
6 Your contribution to statutory health insurance per month	_____ CHF	- 25 CHF
In this situation I choose	☐ the current contract	☐ the new contract

Figure 3.1: Choice question example: Fixed status quo (current contract) vs. alternative (new contract)

2007). Because the intention was to assess interactions and higher-order terms, all possible interactions and higher-order terms up to the power of five were implemented in the design. The 27 choice sets were split randomly into three groups of nine choices each. One choice was included twice in each choice set for consistency checking (Ryan and Bate, 2001), resulting in ten choices per person. In each choice set, the respondents are presented with their (constant) individual status quo and one alternative. Figure 3.1 shows an example. To avoid learning or fatigue effects, the order of the choice alternatives was randomly changed (Kjaer et al., 2006). Some 60 percent of respondents deviated from their status quo at least once. This means that around 40 percent of respondents never chose the alternative insurance contract. In total, 18 percent of the decisions were made in favor of the alternative. As for the consistency test, the choice included twice was "incorrectly" chosen by only 13 of 1,000 respondents. Overall, the observed choices are plausible. Respondents tend to opt for the objectively "good" alternatives and to reject the "bad" alternatives among the ten choices given.

3.3.4 Sample and Interview Strategy

The survey was conducted in Summer 2003 and consisted of 1,000 telephone interviews. Participants were chosen as representative with respect to age, gender, language (the German and the French speaking parts of Switzerland), education, professional status, and rural or urban residence. The survey contained two steps. After individuals agreed to participate, they were asked to look up their personal monthly contributions and their annual deductible to their insurance plan. This guarantees the respondents' knowledge of the status quo, which is essential for making an informed choice between the current contract and a proposed alternative. The participants were also sent an information package, containing descriptions of the attributes. The second step was the DCE itself. Participants were asked to compare the fixed status quo against a hypothetical alternative defined by the attributes mentioned above. The procedure was replicated ten times. Other questions concerned utilization of health care services, preferences for new elements in the insurance package, and socioeconomic characteristics such as age, gender, household income, and education.

Table 3.2: Selected descriptive statistics

Variable	Sample (1)	Official* (2)	Variable	Sample (3)	Official* (4)
N	1,000		Supplementary insurance (%)		
Average age	49	51	- alternative medicine	42	
			- free choice of hospital	39	
Annual deductible (%)			- outpatient treatment	24	
- CHF 230	36	42	- inpatient treatment	21	20
- CHF 400	22	22	(semi-private option)		
- CHF 600	14	10	- insurance coverage abroad	18	
- CHF 1,200	02	03	- inpatient treatment	11	09
- CHF 1,500	25	15	(private option)		
			- dental treatment	11	
Average monthly contribution (CHF)					
	240	280			

Note: * Swiss Population 2003. Source: Federal Office of Public Health (2005); CHF 1 ≈ USD 0.82 at 2005 exchange rates

Table 3.2 shows selected descriptive statistics. Individuals with a low deductible are somewhat underrepresented and individuals with a high deductible are over-represented. On the whole, however, the distribution of the annual deductible is representative when compared with official data (columns (2) and (4)). Since one attribute of interest is coverage of alternative treatment methods, supplementary health insurance is presented as well. 42 percent of the interviewed individuals buy insurance for alternative medicine treatments, making this the most popular supplementary health insurance. Another important dimension is additional coverage of inpatient treatment. 21 percent of individuals have semi-private and 11 percent private accommodation covered by hospital supplementary insurance. In statutory health insurance, basic inpatient services are covered only in hospitals located in the canton of residence. 39 percent of those interviewed buy insurance for free choice of hospitals in all Swiss cantons. The average monthly contributions are CHF 240 in the sample and CHF 280 in official statistics. The discrepancies are explained by three reasons. First, the official statistics include only the (expensive) contracts with the lowest deductible, whereas the sample also includes (less expensive) contracts with higher deductibles and MC alternatives. Second, the canton Ticino, which has traditionally high health care expenditure and also high premiums, is not included in this sample. Third, the official figure includes contributions to accident insurance, which were excluded here.

3.4 Empirical Results

3.4.1 Specification of the Utility Function

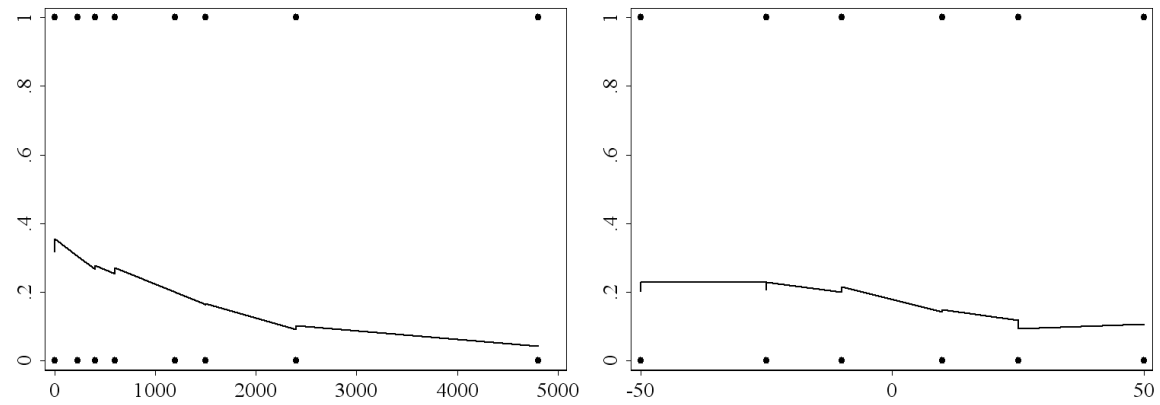
To compare the linear with a non-linear utility function, the non-linearities have to be specified first. Using the data from the DCE presented above, the procedure by Hosmer and Lemeshow (2000) is performed step-by-step as outlined in Section 3.2.2. The results can be summarized as follows.

Step 1: After a univariate analysis of each attribute (not shown here), a forward selection probit estimation with a test of backward elimination is performed. The significance level of entry is set at 0.25 and the level to remove at 0.2. All attributes are approved to belong in the utility function.

Step 2: The multivariate model is estimated and each attribute is then tested according to Step 2 in Section 3.2.2. All attributes prove significance.

Step 3: Method (1): For the continuous variables *Deductible* and *Contribution*, a smoothed univariate scatter plot is estimated (see Figure 3.2). Smoothing is performed with the locally weighted regression command "lowess" in Stata 10.1. Neither attribute bears a linear relation with the dependent variable. The plot for *Deductible* suggests adding a quadratic term. The plot for *Contribution* is non-linear, too. This may call for a cubic term. However, as the result is not conclusive, comparison with results of further steps is required.

Method (2): Figure 3.3 shows the results of the second method to explore the scale of continuous variables by "design variables" (see Step 3 in Section 3.2.2).

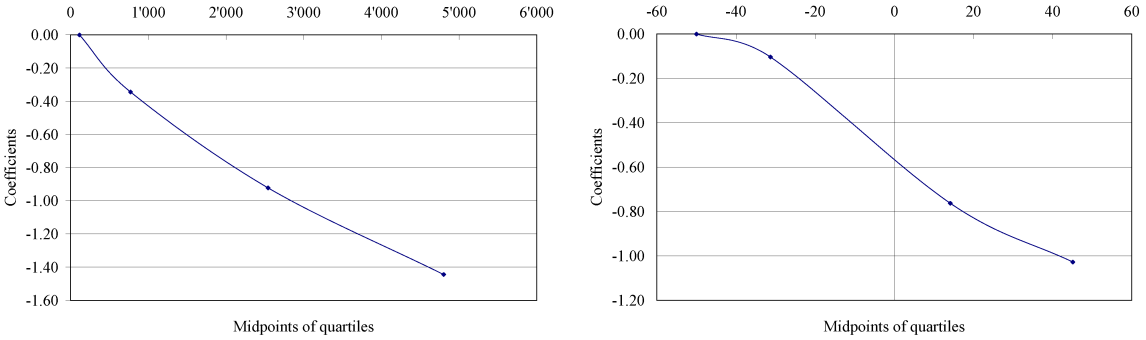


(a) Deductible (b) Health Insurance Contribution

Figure 3.2: Smoothed scatter plots for *Deductible* and *Contribution*

The quartiles' midpoints of *Deductible* and *Contribution*, respectively, are plotted against the estimated probit regression coefficients of the dependent variable (0 if the respondent opted for the status quo, 1 otherwise) on all attributes, with *Deductible* and *Contribution*, respectively, substituted by dummy variables for the quartiles. The plot for *Deductible* shows a linear function with only a slight curvature (see Figure 3.3a). This may indicate a non-linear specification or just a deviation with non-significant implications. The plot for *Contribution* suggests a non-linear, s-shaped relationship (see Figure 3.3b).

Method (3): The third method is the modified Hosmer-Lemeshow test (see Section 3.2.3). For *Deductible*, the test is performed with dummies for 1/8 of the predicted values, and for *Contribution* with dummies for 1/6 of the predicted values. Both F-statistics show that there is a systematic pattern between the residuals and the particular attribute (both p-values of the F-statistics are 0.00). According to these results, *Deductible* and *Contribution* should not be specified as linear.



(a) Deductible (b) Health Insurance Contribution

Figure 3.3: Plots of estimated probit regression coefficients vs. approximate quartile midpoints of *Deductible* and *Contribution*

Note: Coefficients are from a regression of the dependent variable (0 if the respondent opted for the status quo, 1 if he or she opted for the alternative) on three dummies for the second, third, and fourth quartiles of *Deductible* and *Contribution*, respectively, and the remaining attributes. In the plot the coefficient of the first quartile's dummy is set to zero.

Step 4: To test for interactions, the LR test is used. Every possible interaction between the attributes is tested. For evidence about the functional form of the attributes, the LR test is also performed for terms of second and higher orders. The interactions proven to be significant are presented in Table 3.3.

Summarizing the findings from Steps 3 and 4, the *Deductible* smoothed scatter plot suggests a squared specification (see Figure 3.2a). Figure 3.3a favors a linear or quadratic form. Whereas these methods do not draw a final conclusion (the deviations could be non-significant implications for the specification), the modified Hosmer-Lemeshow test clearly favors a non-linear utility function. The LR tests in Step 4 support this finding. According to the LR test, both a squared and a cubic term significantly contribute to an improvement of fit. Backward selection procedures and goodness-of-fit tests are required to decide the final specification.

The smoothed scatter plot for *Contribution* (see Figure 3.2b) and the "design variables" (see Figure 3.3b) suggest a non-linear specification. The Hosmer-Lemeshow test confirms this result. However, it is unclear how many higher-order terms should be included. The LR tests suggest going to the fourth power. Backward selection and goodness-of-fit tests are also required for the final specification of this attribute.

Step 5: A backward selection procedure is performed and the results are assessed in view of goodness-of-fit. The most parsimonious utility function that still passes all specification tests is the following: Besides all main effects, interac-

Table 3.3: Interactions and terms of higher order resulting from Step 4

$Copayment \times Alternative$	$Wait\_Innovation \times Generics$
$Copayment \times Generics$	$Wait\_Innovation \times Contribution$
$Copayment \times Contribution$	$Alternative \times Contribution$
$Contribution^2$	$Deductible^2$
$Contribution^3$	$Deductible^3$
$Contribution^4$	

Note: For labels see Table 3.1

Table 3.4: Comparison of goodness-of-fit results linear and non-linear utility function

Test statistic	Linear utility function (1)	Non-linear utility function (2)
Link test p-value	0.00	**0.15**
Reset test F-stat p-value	0.00	**0.09**
Copas test F-stat p-value	**0.90**	0.71
Modified HL test F-stat p-value	0.00	**0.80**
HL test χ^2-stat p-value	0.00	**0.63**
Pearson Correlation Coefficient	0.03	**0.00**
- p-value	0.00	**0.74**
AIC	7,272	**6,320**
BIC	7,322	**6,442**
LL	−3,629	**-3,143**
LR test		0.00
N	9,655	9,655

Note: AIC is the Akaike Information Criterion and BIC is the Bayesian Information Criterion. LL is the log likelihood, evaluated at the maximum likelihood estimator. Boldface entries indicate the better specification for the particular criterion.

tions are *Alternative* × *Copayment*, *Wait_Innovation* × *Generics*, *Copayment* × *Generics*, *Wait_Innovation* × *Contribution*, and *Copayment* × *Contribution* (note that *Alternative* × *Contribution* is dropped compared to Table 3.3). *Deductible* has to be included in squares and *Contribution* to the power of four. This will later be referred to as the "non-linear" utility function, as opposed to the "linear" utility function containing the main effects only.

3.4.2 Comparison of Goodness-of-Fit

In this section, the utility functions are compared with regard to goodness-of-fit. The results of the tests presented in Section 3.2.3 are shown in Table 3.4. The linear specification fails in all but one test (see column (1)). Neither the Link, the Reset, nor one of the Hosmer-Lemeshow tests support the linear utility function. This specification only passes the Copas test. However, this was expected, since Copas is a test for over-fitting the data and the model is very parsimonious, containing only the main effects. The non-linear utility function passes all the tests presented. Only the Reset test might attract attention, with a p-value of 0.09. Strict adherence to a

10 percent level of statistical significance would point out a possible s-shaped misfit. However, considering the richness of the data set (1,000 respondents with ten decisions each, resulting in 10,000 observations), it is important not to rely too strictly on test statistics. The p-value most likely does not point out a significant misfit in this case.

Performing the modified Hosmer-Lemeshow test, the size of the data set allows the building of 40 groups, each comprising 2.5 percent of the predicted values. Figures 3.4 and 3.5 present the coefficients from regressing the residuals (from a regression of the dependent variable on the attributes' main effects (Figure 3.4) or the main effects and the additional variables for non-linearity (Figure 3.5), respectively) on dummies for the 40 groups. If there was no misspecification, the coefficients would be distributed randomly among the zero-line. However, Figure 3.4 shows a systematic u-shaped pattern. Consequently, the linear specification fails the test of no misfit. The pattern for the non-linear utility function (see Figure 3.5) passes the test with a p-value of 0.80 when testing the residuals to be zero jointly. From Figure 3.5, the coefficients are seen to be distributed close to randomly among the zero-line. Both the Akaike and Bayesian information criteria prefer the non-linear utility function, as well as the log-likelihood. An LR test shows that the added variables significantly contribute to a better fit of the utility function.

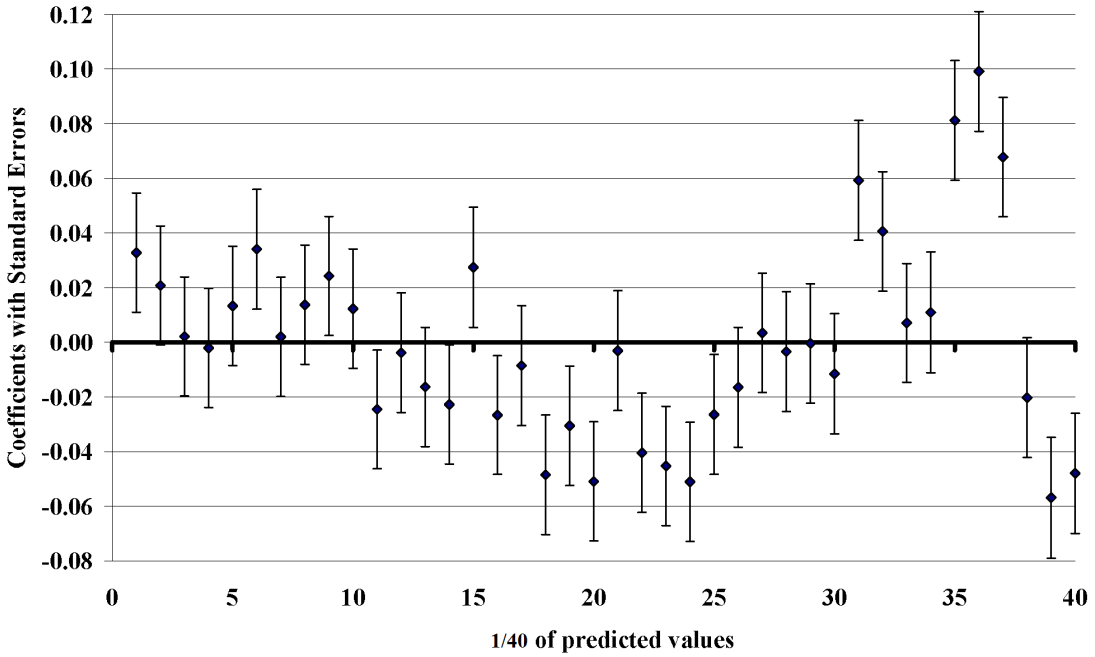


Figure 3.4: Regression result modified Hosmer-Lemeshow test linear utility function

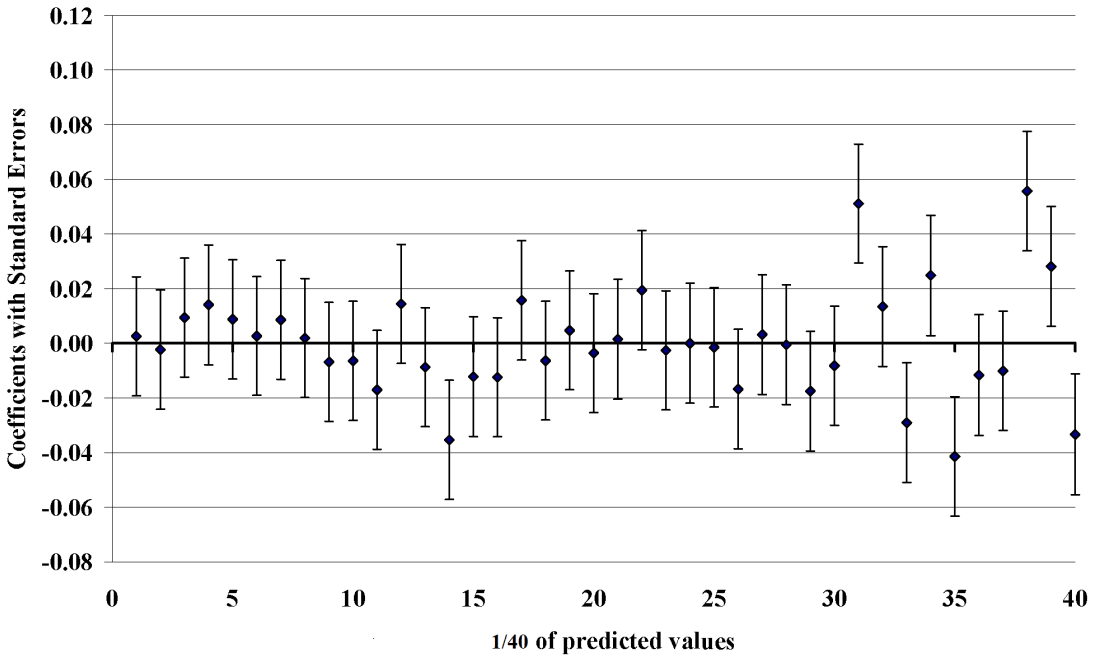


Figure 3.5: Regression result modified Hosmer-Lemeshow test non-linear utility function

Note: Coefficients are from a regression of the residuals on dummy variables for 1/40 of the predicted values.

3.4.3 Comparison of Willingness-to-Pay

In this section, the utility functions are compared with regard to estimated WTP. If there is no significant difference, the linear utility function can be seen as a good approximation. If the results differ significantly, it emphasizes the necessity to apply statistical specification rules in DCE analysis.

Table 3.5 shows the estimated WTP values for the linear utility function (WTP_l, column (1)) and the non-linear utility function (WTP_n, column (4)).[5] All WTP are calculated in terms of health insurance contributions, i.e. β_n in Equation (3.5) is the coefficient of the attribute *Contribution*. Because WTP is not constant with the non-linear utility function, the estimated values are stated subject to the other attributes being equal to the status-quo level. Columns (2) and (5) of Table 3.5 show standard errors according to the delta method. As Mullahy and Manning (1996) note in the context of cost-effectiveness analysis, the delta method does not work well in the case of ratios. They present two safe strategies for calculating confidence intervals of ratios, one of which is bootstrapping. For this purpose, both utility functions have been bootstrapped simultaneously with 1,000 iterations (see columns (3) and (6) for the standard errors). This approach allows comparison of the estimates directly, t-testing the bootstrapped differences in WTP. This result is shown in column (8) of Table 3.5.

With both utility functions, the estimated WTP values are significantly different from zero. However, the difference between WTP_l and WTP_n is of considerable magnitude and is statistically significant in three out of five attributes. (1) The linear approach suggests that the respondents are willing to pay CHF 24 per month for the coverage of more alternative treatment methods (*Alternative*). With the non-linear specification, this value decreases to CHF 12 per month. (2) For a more restrictive reimbursement

[5]For the estimated coefficients see Table 3.6 in the Appendix. The coefficients of the higher-order terms are rather small, but nevertheless significantly different from zero. A specification neglecting these terms was tested for the sake of parsimony. However, in this case the Reset test signals a clear misfit, as does the (modified) Hosmer-Lemeshow test.

of pharmaceuticals (*Generics*), respondents have to be compensated with a reduction in health insurance contributions of CHF 13 per month when estimating the linear specification. This willingness-to-accept more than doubles to CHF 29 when using the non-linear function. (3) Concerning financing, respondents ask for a reduction of CHF 19 in health insurance contributions (WTP_l) with an increase in the copayment rate (*Copayment*) from 10 to 20 percent. With the non-linear specification, respondents ask for a reduction of CHF 32 per month. However, the differences in compensation for delayed coverage of treatment methods (*Wait_Innovation*), a marginal increase in the deductible (*Deductible*), and the difference in the status-quo bias (*Constant*) are small and non-significant.

The specification of the utility function has considerable effects on estimated WTP. The consequences of using the linear specification as a simplification can be striking. In the case of the presented DCE, these might include the following. (1) A health insurer might launch an alternative contract reimbursing more alternative treatment methods for an increase in premiums of CHF 24 per month (assuming this amount covers costs). The number of people actually buying the contract will be much lower than expected, because enrollees are willing to pay, on average, an increase of only CHF 12 per month for the additional benefits. (2) The regulator may propose to reimburse only the cheapest pharmaceutical product on the list of benefits and to decrease health insurance premiums by CHF 13 per month. However, the decrease in contributions is lower than the amount people ask for cuts in benefits. Inefficiencies result. (3) One may propose to decrease health insurance contributions by CHF 19 per month in exchange for an increase in the copayment rate from 10 to 20 percent and a maximum of CHF 1,200 instead of CHF 600 per year, with the advantage of mitigating moral hazard. However, the results show that respondents ask for higher compensation (CHF 32 per month) to accept this increase. This proposition might thus cause inefficiencies.

Table 3.5: Marginal WTP linear and non-linear utility function, CHF per month

	Linear Specification			Non-linear Specification			Difference	t-statistic†
		Standard Errors			Standard Errors			
Attribute	WTP_l	Delta Method	Bootstrap	WTP_n	Delta Method	Bootstrap		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Constant	-31.87***	4.55	4.92	-30.52***	11.71	13.94	-1.35	0.13
Alternative[1]	23.89***	3.09	3.46	11.70***	5.28	5.58	12.19	2.52
Generics[2]	-12.47***	3.03	3.26	-29.03***	7.28	7.35	16.56	2.69
Wait_Innovation[3]	-39.40***	3.34	3.79	-38.11***	5.89	6.25	-1.29	0.15
Copayment[4]	-19.08***	3.00	3.38	-31.71***	4.49	4.43	12.63	3.02
Deductible [5]	-00.03***	0.00	0.00	-00.04***	0.01	0.01	0.01	1.51

[1] Reimbursement of more alternative treatment methods (alternative practitioner and naturopathy)

[2] Reimbursement of only the cheapest pharmaceutical

[3] Coverage only two years after approval

[4] 20 percent with a max. of CHF 1,200 / Year

[5] Choice between CHF 0, 2,400, and 4,800 / Year

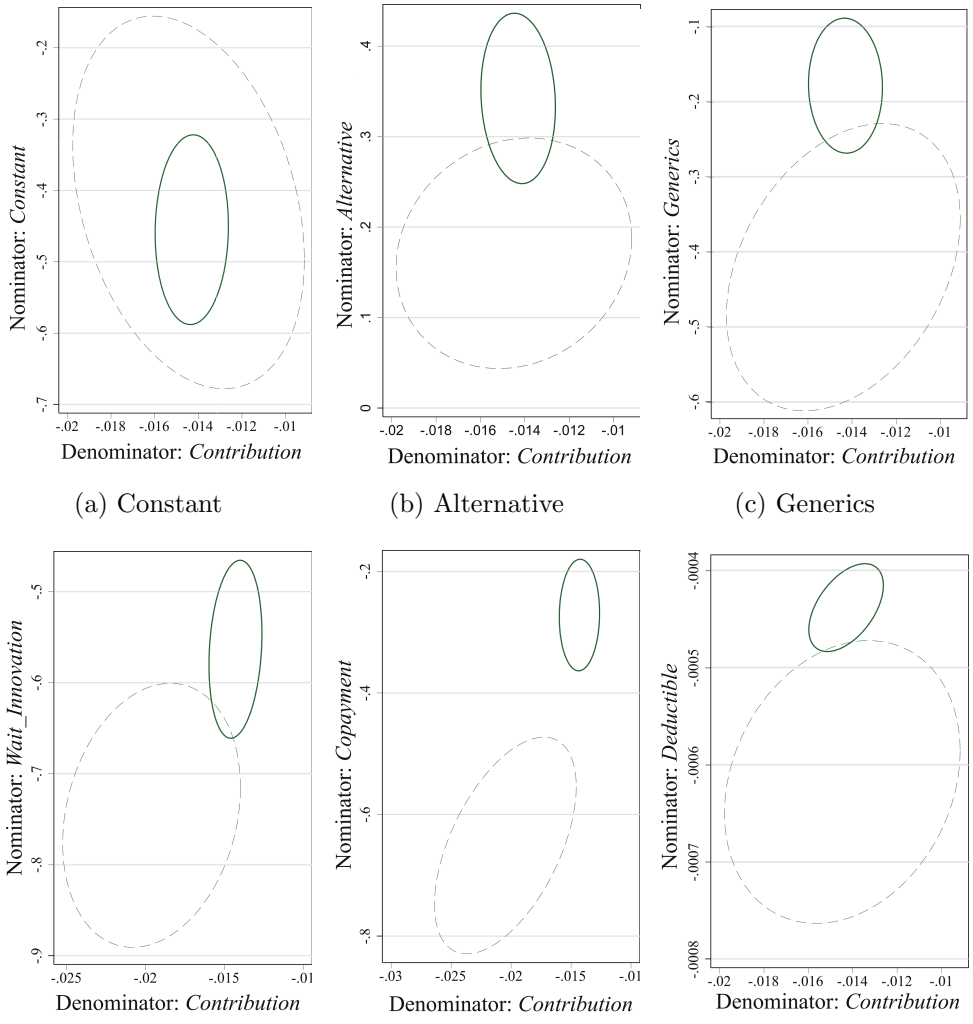
Note: *** indicates significance at the 1 percent level; †: t-test statistics from one bootstrap with both models estimated simultaneously, nr. of iterations = 1,000, no bias correction required (see Efron and Tibshirani, 1993), no acceleration; CHF 1 ≈ USD 0.82 at 2005 exchange rates

Simultaneously bootstrapping the linear and non-linear utility functions allows the comparison of the two sets of WTP values. Figure 3.6 shows 95 percent confidence ellipses,[6] where the solid line corresponds to the linear function and the dashed line to the non-linear utility function. The x-axes are the denominators of the marginal WTP (where the marginal WTP is equal to the marginal rate of substitution, see Equation (3.5)). The y-axes are the numerators. In the case of the linear utility function, the axes equal the single estimated coefficients (the x-axis is always equal to *Contribution*, the y-axes vary). In the case of the non-linear utility function, the x-axes are single coefficients (for *Constant*, *Alternative*, *Generics*) or combinations of coefficients (for *Wait_Innovation*, *Copayment*, and *Deductible*). The y-axes are combinations of coefficients for all attributes in this case.

From Table 3.5, it can be seen that the standard errors become larger with the non-linearities. This increase is visible in Figure 3.6 as well, where the ellipses become larger. The positions of the pairwise ellipses provide information about the origin of differences in WTP estimates. For example, in the plot for the coverage of alternative treatment methods (Figure 3.6b), the origins of the two ellipses have almost the same x-coordinate, but different y-coordinates. The decrease in WTP from CHF 24 (WTP_l, see Table 3.5) to CHF 12 (WTP_n) therefore mostly results from the change in *Alternative*. Hence, respondents value a marginal increase in *Contribution* almost the same. However, coverage of additional treatment methods is valued less with the non-linear specification than with the linear. The same holds for *Generics* (Figure 3.6c), where the increase in willingness-to-accept from CHF 13 to CHF 29 per month results

[6]The estimated coefficients are normally distributed by assumption. Linear combinations of normal random variables are normally distributed as well, and so Figure 3.6 shows bivariate normal distributions. These belong to the elliptical family (McNiel et al., 2005). This makes the confidence curves ellipses. The ellipses can also be interpreted as visual indicators of correlation (SAS Institute Corp., 1999). An ellipse collapses diagonally as the attributes become perfectly positively or negatively correlated. In case of uncorrelated attributes, the ellipse is circular or the orientation is aligned with a coordinate axis. If the ellipse's orientation is to the northeast or the southwest quadrant, the attributes are positively correlated. If it is towards the northwest or the southeast quadrant, there is negative correlation. Further, from the orientation of the ellipses, it can be seen whether the 95 percent confidence regions are likely to include the origin or not. If this were the case, then the delta-method approximation would fail (see Gleser and Hwang, 1987). The Fieller method would provide the appropriate alternative (see Fieller, 1954, and for applications e.g. Willen and O'Brien, 1996 or Heitjan, 2000).

mostly from *Generics*. In the case of *Copayment* (Figure 3.6e), the valuation of both attributes changes with the specification. The x- and the y-coordinates are shifted. This is due to the interaction term $Copayment \times Contributions$.



(a) Constant (b) Alternative (c) Generics

(d) Innovation (e) Copayment (f) Deductible

Figure 3.6: 95 percent confidence ellipses for marginal WTP

Note: Solid line = linear utility function, (de)nominator is equal to estimated coefficient (see Equation 3.4); dashed line = non-linear utility function, (de)nominator is equal to estimated coefficient (nominator: *Constant*, *Alternative*, *Generics*) or linear combination of coefficients (nominator: *Wait_Innovation*, *Copayment*, *Deductible*; denominator: *Contributions*).

3.5 Conclusions

When estimating willingness-to-pay (WTP) using discrete-choice experiments (DCEs), the utility function is most commonly assumed to be linear in the attributes. Interactions and higher-order terms are set to zero with reference to Louviere et al. (2000). Non-linearities are only implemented if they are of special interest or in the guise of interactions with socioeconomic characteristics to investigate heterogeneity in preferences. However, it can be argued that the utility function is unlikely to be linear because of diminishing marginal utility and gain-loss asymmetries (Hoyos, 2010). This paper addresses the issue of the utility function's form by showing that the linear approximation can be a risky choice in DCEs. For this purpose, an experiment conducted by Becker (2006, Chapters 6-8) in Switzerland is reexamined. The DCE assesses preferences for Swiss statutory health insurance and WTP for proposed reforms. The attributes describe changes in the list of benefits (inclusion of additional alternative treatment methods, reimbursement of only the cheapest pharmaceuticals (generics), delayed access to new treatment methods) and changes in financing (increased copayment rate, change in deductible and health insurance contributions).

The DCE is estimated in two ways, first using a linear utility function, including the main effects only, and second using a non-linear utility function, allowing for interactions and terms of higher order. The procedure of Hosmer and Lemeshow (2000) is used for the specification of non-linearities, following the statistical model-specification literature. In a first step, the utility functions are tested for misfits, using a variety of goodness-of-fit measures as proposed by Basu et al. (2004) and Basu et al. (2006). The linear utility function is found to perform better with regard to over-fitting, but to have serious misfit problems. However, the non-linear utility function is found to present the data well. In a second step, the utility functions are compared with regard to estimated WTP, stated in terms of health insurance contributions. The results are found to significantly differ in terms of statistical significance and in magnitude for three out of five attributes. (1) With the linear utility function, respondents are willing to pay CHF 24 per month for the reimbursement of additional alternative

treatment methods. With the non-linear specification, estimated WTP decreases to CHF 12 per month. (2) The linear specification proposes that respondents must be compensated with a decrease in contributions of CHF 12 per month in order to accept the reimbursement of only the cheapest pharmaceuticals (generics). However, this willingness-to-accept more than doubles with the non-linear utility function at CHF 29 per month. (3) An increase in the copayment rate from 10 to 20 percent with a simultaneous increase in the maximum copayment from CHF 600 to CHF 1,200 per year must be compensated with a decrease in contributions of CHF 19 per month with the linear specification, but rises to CHF 32 per month with the non-linear specification.

These findings suggest that the form of the utility function can have significant impact on estimated WTP. Using the linear specification as an approximation may lead to seriously biased estimates. Since DCEs are playing an increasingly significant role in health care decision making, the assumption of a linear utility function may lead to inefficient use of health care resources.

However, this research is subject to several limitations. It can be argued that the linear functional form is a poor approximation in this setting, but might sufficiently serve in others. Also, there is no "best" way of finding an appropriate model specification. The methodological debate is still ongoing, and will certainly continue into the future. A disadvantage of the procedure presented here is that finding the non-linearities is very time-consuming. The limiting factor is not computer power but rather creativity in finding a model that passes all specification tests simultaneously. Finally, the form of the utility function is one among many aspects of DCEs that need further consideration. Issues range from questionnaire development and choosing experimental designs to applications of new econometric methods (for a summary see Louviere and Lancsar, 2009 or Hoyos, 2010). However, given the presented findings, the form of the utility function is an important issue in DCE analysis which should be taken into account for future DCEs.

Acknowledgments

The author gratefully acknowledges econometric advice by Willard G. Manning (University of Chicago, Illinois, USA) and helpful suggestions by Peter Zweifel (University of Zurich, Switzerland).

Appendix

Table 3.6: Results random-effects probit estimation, linear and non-linear utility function

	Linear Utility Function		Non-linear Utility Function	
Variable	Coefficient	St.error	Coefficient	St.error
Constant	-0.4548^{***}	0.0617	-0.4191^{***}	0.1332
Alternative	0.3409^{***}	0.0419	0.1692^{***}	0.0626
Generics	-0.1779^{***}	0.0417	-0.4198^{***}	0.0924
Wait_Innovation	-0.5623^{***}	0.0480	-0.7377^{***}	0.0735
Copayment	-0.2723^{***}	0.0428	-0.6424^{***}	0.0924
Deductible	-0.0004^{***}	0.0000	-0.0006^{***}	0.0001
Contribution	-0.0143^{***}	0.0006	-0.0157^{***}	0.0026
Copayment × Alternative			0.3311^{***}	0.0940
Wait_Innovation × Generics			0.3808^{***}	0.1177
Copayment × Generics			0.2058^{*}	0.1090
Wait_Innovation × Contribution			-0.0050^{***}	0.0015
Copayment × Contribution			-0.0060^{***}	0.0015
Deductible^2			$4.e-08^{***}$	$1.e-08$
Contribution^2			0.0006^{**}	0.0003
Contribution^3			$2.e-06^{**}$	$1.e-06$
Contribution^4			$-2.e-07^{**}$	$9.e-08$
σ_μ	0.9639^{***}	0.0426	0.9941^{***}	0.4412
ρ	0.4816^{***}	0.0221	0.4970^{***}	0.0222

Note: *** indicates significance at the 1, ** at the 5, and * at the 10 percent level

Chapter IV
Capping Risk Adjustment?

Patrick Eugster, Michèle Sennhauser
and Peter Zweifel

Published in

Journal of Health Economics,
Vol. 29 (4), p. 499-507, 2010

4 Capping Risk Adjustment?

4.1 Introduction

When premiums are community-rated, risk adjustment (RA) is introduced in order to reduce competitive insurers' incentive to select favorable risks. On the other hand, unless fully prospective, it undermines their incentive for efficiency (Ellis and Van de Ven, 2000). This goal conflict has been addressed by Van Barneveld et al. (2001), who estimate optimal thresholds in the cost distribution beyond which RA sets in. However, the implementation of such a rule becomes difficult when the distribution of health care expenditure (HCE) not only shifts over time (e.g. due to a particularly high rate of cost increase in the hospital sector) but also differs between insurers (e.g. due to more or less reliance on Managed-Care (MC) options). Therefore, as a rough-and-ready measure, one might consider simply capping the volume of RA. In this way, one exposes insurers to some residual financial risk. At the same time, there is dissatisfaction with the performance of current RA formulas (Van de Ven et al., 2003). The expectation is that by adding risk adjusters, incentives for risk selection could be reduced even more. However, refinements of the RA formula quite likely cause the volume of RA to increase. They therefore are in conflict with the desire to preserve efficiency through capping the volume of RA.

There is a second motivation for limiting the volume of RA, which becomes evident as soon as one recognizes the analogy between RA and a levy (and a subsidy, respectively). In the economics of taxation, a distinction is consistently made between those who pay a levy and those who ultimately bear it. In the case of RA, payments

into the scheme are ultimately borne by the favorable risks whose premiums exceed the (estimated) actuarially fair value. They in fact cross-subsidize the premiums of unfavorable risks. A part of this cross-subsidization occurs through community-rated premiums. The remainder is paid (but not borne) by health insurers. Analysis of RA schemes so far has exclusively focused on this second, more visible component. It is designed to neutralize insurers' incentives to select favorable risks and is often referred to as "volume of RA". To avoid confusion, this definition will be used in this chapter[1]. However, note that it is the total amount of cross-subsidization that drives consumer behavior. Favorable risks have an interest in avoiding the cross-subsidy by seeking out an insurer who offers a premium closer to the actuarially fair volume, which remains possible as long as RA is not perfect. Conversely, unfavorable risks have an interest in obtaining a high cross-subsidy through their choice of insurer.

Both concerns have become important in Switzerland, a country with competitive social health insurance. By 2005, *cross-subsidization (CS, between individuals)* had reached CHF 4.8 billion (bn.), about 1 percent of Swiss GDP (CHF 1 $\approx$ USD 0.83). Thus, RA had turned into a redistributive scheme in its own right. The *volume of risk adjustment (RA, between insurers, partly retrospective)* amounted to CHF 1.2 bn., or some 20 percent of their payments for HCE. This was seen as incompatible with the efficiency goal by the Swiss Council of States, who considered limiting the volume of RA to its 2004 value (inspired by the capping-proposal by Spycher, 2000). At the same time, parliament decided to add the criterion, "Hospitalization or living in a nursing home during the previous year" ("hospitalization" henceforth) to the RA formula, effective 2012. This decision was influenced by Beck et al. (2006), who had found that this criterion serves to substantially reduce insurers' payoff to risk selection.

With this backdrop, the present contribution purports to achieve two objectives. First, it seeks to establish the opportunity cost of capping the volume of RA in terms of increased incentives for risk selection. Second, it investigates the consequences of

[1]Interestingly however, Swiss statistics also do publish the amount of cross-subsidization between individuals as "volume of RA" (see Joint Organization KVG, 2005).

complementing the RA formula by the criterion "hospitalization". The main results of this paper are twofold. First, the introduction of the hospitalization adjuster is shown to inflate the volume of CS and of RA in every canton of Switzerland, in some of them by more than 50 percent. Second, reducing the amount of CS from an estimated CHF 5.375 bn. to CHF 4.5 bn. (its 2004 volume) would reduce the volume of RA between insurers from an estimated CHF 1.1 bn. to CHF 0.98 bn. To minimize the associated incentives for insurers to increase their risk-selection efforts, the new RA values have to be the higher, the greater the differences between group-specific values prior to the limitation and the greater the group's population share.

This chapter is organized as follows. Section 4.2 is devoted to a description of the data basis, descriptive statistics, and checks of the representativeness of the data. Section 4.3 shows how the volumes of CS and RA change when the additional criterion is introduced. Section 4.4 then takes up the issue of capping the volume of risk adjustment. The consequences for consumers, insurers, and Swiss cantons are analyzed. In Section 4.5, limiting the volume of RA is accepted as a way to preserve insurers' incentives for efficiency, giving rise to an optimization problem since insurers' tendency to turn to risk selection again should be minimized. Section 4.6 discusses the consequences of capping RA, illustrating them with an empirical example. Section 4.7 concludes.

4.2 Data Basis, Descriptive Statistics, and Representativeness

In order to carry out this research, three large Swiss health insurers provided individual health insurance data. Their holders of basic health insurance during the period 2001 to 2005 were considered, totaling 2.78 million (mn.) individuals. Besides socioeconomic variables like age, gender, and canton of residence, data on ambulatory and hospital health care expenditure (HCE), drug expenditure, and a variable indicating hospitalization or living in a nursing home in the previous year were collected. To characterize the type of health insurance, the deductible and a

variable indicating choice between conventional and MC contracts were included as well. With 49.5 percent of women, the sample is well balanced with respect to gender. The market share of the three insurers is stable across the age profile, amounting to 25 percent on average. Across the 26 Swiss cantons, they are over-represented in eastern and central Switzerland and under-represented in the northern and western parts of the country.

In Swiss health insurance, premiums are community-rated. They are uniform in 16 cantons, the remaining cantons distinguishing up to three premium regions. In 2005, 32 percent of the population lived in cantons with uniform premiums, while 25 percent lived in a high, 27 percent in a medium, and 16 percent in a low premium region. With regard to the choice of contract, there is a clear trend toward higher deductibles. Whereas in 2001, policies with the minimum deductible (which amounts to CHF 300 or USD 250 at 2007 exchange rates) had a share of over 56 percent, this share had decreased to 50 percent by 2005. The three highest deductibles (CHF 1,500, 2,000, and 2,500, respectively) increased in importance from 12 to over 22 percent during the same period. There is a similar trend in favor of MC contracts. Especially consumers aged 31 to 35 use this option, resulting in a share of 22 percent in this age group. However, older age cohorts increasingly prefer MC contracts as well. For instance, among the over 80 year old, their share went up from 10 to 18 percent between 2001 and 2005.

In Swiss RA, only two criteria are considered, age and gender. The age classification comprises 15 classes, starting from 19-25 and continuing in 5-year steps. By law, RA must not lead to financial reallocation between cantons. The national volumes of CS and RA therefore equal the sum of the cantonal volumes. Computing these volumes using the sample data of the three insurers and their market shares yields a CS total of CHF 4.13 bn., the national figure being CHF 4.8 bn. (see Joint Organization KVG, 2005).

Table 4.1: Volume of cross-subsidy per canton, CHF (2005)

Canton	mn.*	mn.**	p.c.*	p.c.**	Canton	mn.*	mn.**	p.c.*	p.c.**
ZH	735.82	796.47	713	771	SH	39.18	47.50	649	787
BE	592.03	678.17	768	879	AR	20.10	24.85	483	597
LU	193.80	211.23	691	753	AI	6.48	7.43	570	654
UR	17.57	19.97	636	723	SG	196.02	238.19	545	662
SZ	57.00	70.85	533	662	GR	95.91	110.26	616	708
OW	13.01	16.41	497	626	AG	259.61	312.05	573	689
NW	15.13	17.23	485	552	TH	105.30	126.78	579	697
GL	16.52	19.89	544	656	TI	226.03	270.02	869	1,038
ZG	47.57	54.14	566	644	VD	430.40	502.39	852	994
FR	121.48	146.44	625	754	VS	160.39	170.13	685	726
SO	122.07	154.65	614	778	NE	104.10	130.63	784	984
BS	110.21	161.97	719	1,057	GE	257.43	333.04	816	1,056
BL	149.81	172.14	697	800	JU	39.45	52.32	733	972

Note: * Simulation, ** Official data from Joint Organization KVG (2005), p.c. = per capita, CHF 1 $\approx$ USD 0.83 at 2007 exchange rates

In all cantons, the calculated volumes of CS (see Table 4.1) and RA fall short of the official ones. The difference is smallest in the canton of Zurich (ZH), Lucerne (LU), and Valais (VS), amounting to less than 10 percent. It is between 10 and 20 percent in 9 other cantons, where the three insurers only hold small market shares. This marked discrepancy could reflect successful risk-selection efforts, which have high expected return if targeted at a small population at risk (as shown in Zweifel and Eisen, 2005, Chapter 5.5). However, there is no significant (negative) correlation between market shares of the three insurers and deviations from the official CS and RA figures, suggesting that risk selection is not the explanation.

Table 4.2 focuses on CS values, in accordance with the argument proffered in the Introduction that they are the ones that trigger risk-selection effort on the part of consumers. Calculated cross-subsidies per capita for all 30 groups used in RA are shown, along with their standard errors and official countrywide values. Young men have to bear the highest cross-subsidies (over CHF 2,000 per year), followed by young women with CHF 1,773 per year. Over 90 year old women benefit the most, to the tune of over CHF 8,600, followed by women of age 86 to 90 with CHF 6,917 and

men of age 90+ with CHF 6,731. All age groups over 60 are cross-subsidized by the combination of community-rating and RA.

Table 4.2: Simulated and official cross-subsidies per capita according to age and gender, CHF (2005)

Men	Average*	Std.error	Min	Max	Official value
19-25	−2,006.50	505.52	−3,006.17	−707.84	−1,963.87
26-30	−1,227.59	833.80	−2,165.91	2,287.40	−1,889.64
31-35	−900.68	678.91	−1,733.38	1,202.03	−1,771.42
36-40	−979.03	421.93	−1,749.27	247.62	−1,624.49
41-45	−828.69	351.55	−1,435.17	−40.31	−1,398.94
46-50	−543.46	465.97	−1,615.88	349.08	−1,091.94
51-55	−109.82	378.55	−977.63	714.71	−624.63
56-60	290.34	300.27	−557.57	815.53	13.40
61-65	884.74	418.34	228.53	1,648.89	771.06
66-70	1,560.60	598.50	187.69	2,464.57	1,638.40
71-75	2,535.19	548.54	982.57	3,435.54	2,873.43
76-80	3,208.98	653.35	1,884.58	4,128.30	3,845.50
81-85	4,127.79	1,361.80	1,261.52	6,983.73	4,986.30
86-90	5,286.51	1,208.24	2,752.09	7,945.75	6,880.09
90+	6,731.78	1,513.63	2,945.10	8,915.78	9,541.96
Women	Average*	Std.error	Min	Max	Official value
19-25	−1,772.99	494.20	−2,780.08	−974.44	−1,484.37
26-30	−1,024.61	461.54	−2,211.50	−311.71	−946.01
31-35	−746.06	559.49	−1,694.31	−1,125.73	−749.83
36-40	−961.00	328.45	−1,576.69	−316.11	−924.81
41-45	−965.85	279.05	−1,749.34	−535.99	−922.02
46-50	−732.01	309.04	−1,295.60	−177.44	−646.82
51-55	−442.87	268.14	−1,045.08	106.95	−235.80
56-60	−15.51	321.10	−512.16	841.85	205.36
61-65	443.65	247.14	19.55	764.95	737.31
66-70	981.80	395.53	210.13	1,603.77	1,415.39
71-75	1,982.76	446.04	758.34	2,662.32	2,385.07
76-80	3,136.84	656.22	1,838.10	4,406.12	3,671.81
81-85	4,641.23	775.55	2,788.30	6,111.25	5,596.14
86-90	6,917.12	987.66	5,115.11	8,382.98	8,486.06
90+	8,672.75	1,770.15	4,464.86	11,619.96	12,457.28

Note: * Average over all 26 Swiss cantons, CHF 1 ≈ USD 0.83 at 2007 exchange rates

A comparison with official values (see the last column of Table 4.2) shows calculated values to be too high for younger and too low for older individuals, especially for women. These deviations are mainly responsible for the underestimation of the total CS and RA volumes noted above. Table 4.2 also shows that the variance of CS values increases with age. While the standard deviation is CHF 494 for young women, it attains CHF 1,770 for the oldest age class, reflecting the fact that the variance of HCE increases with age. Overall, calculated figures come close enough to official CS values to justify the use of sample data in the investigation below.

4.3 Hospitalization as an Additional Criterion

Current Swiss RA uses only the two criteria age and gender. However, the hospitalization adjuster will be added to the RA formula from 2012. Beck (2004) and Beck et al. (2006) estimate that this criterion has considerable predictive power in explaining future HCE. To prevent gaming by insurers, stays of less than four days are not counted. Maternity stays are excluded as well since according to Beck (2004), they do not significantly increase HCE in the following year. The new RA formula will continue to be partly retrospective because it uses observed rather than predicted HCE values.

The new criterion has several advantages. It is very easily collected; moreover, being a dummy variable it does not make computation of RA payments much more difficult. While the formula currently distinguishes 30 age-gender cells, the number of classes would only increase to 60 (for a discussion on other alternatives and their drawbacks see e.g. Lamers, 1999, Ellis and Van de Ven, 2000, Lamers and Van Vliet, 2003a, Lamers and Van Vliet, 2003b, Van de Ven et al., 2004, Beck et al., 2006, and Van de Ven et al., 2007). Moreover, the data is readily available in every insurer's administrative database.

Taking this additional criterion into account, calculated cross-subsidies would increase from CHF 4.13 bn. (as of 2005) to CHF 5.82 bn., i.e. by 40 percent. According to

Table 4.3, every canton would exhibit an increase. To put these into perspective, note that premiums e.g. in the canton of Zurich were CHF 4,000. Therefore, the per-capita cross-subsidy of CHF 870 would have attained almost 22 percent of premium under the new RA formula.

Moreover, the change would have completely overthrown the customary CS age and gender profiles. Whereas under the current RA formula, the young of both genders are

Table 4.3: Cross-subsidization without and with the hospitalization adjuster, CHF (2005)

	Without the criterion		With the criterion		
Canton	CHF mn.	CHF per capita	CHF mn.	CHF per capita	% Increase
ZH	735.82	713	898.24	870	22.1
BE	592.03	768	833.26	1,080	40.7
LU	193.80	691	292.31	1,042	50.8
UR	17.57	636	26.00	941	48.0
SZ	57.00	533	84.29	788	47.9
OW	13.37	497	22.98	877	71.9
NW	15.13	485	21.35	684	41.0
GL	16.52	544	24.23	798	46.6
ZG	47.57	566	71.94	855	51.2
FR	121.48	625	184.14	948	51.6
SO	122.07	614	176.52	889	44.6
BS	110.21	719	171.37	1,118	55.5
BL	149.81	697	201.91	939	34.8
SH	39.18	649	59.72	989	52.4
AR	20.10	483	32.94	791	63.9
AI	6.48	570	9.58	842	47.8
SG	196.02	545	298.32	829	52.2
GR	95.91	616	137.59	884	43.5
AG	259.61	573	363.64	803	40.1
TH	105.30	579	173.73	955	65.0
TI	226.03	869	313.01	1,203	38.5
VD	430.40	852	593.60	1,175	37.9
VS	160.39	685	226.85	969	41.4
NE	104.10	784	166.65	1,255	60.1
GE	257.43	816	375.23	1,190	45.8
JU	39.45	733	54.38	1,010	37.8

Note: CHF 1 $\approx$ USD 0.83 at 2007 exchange rates

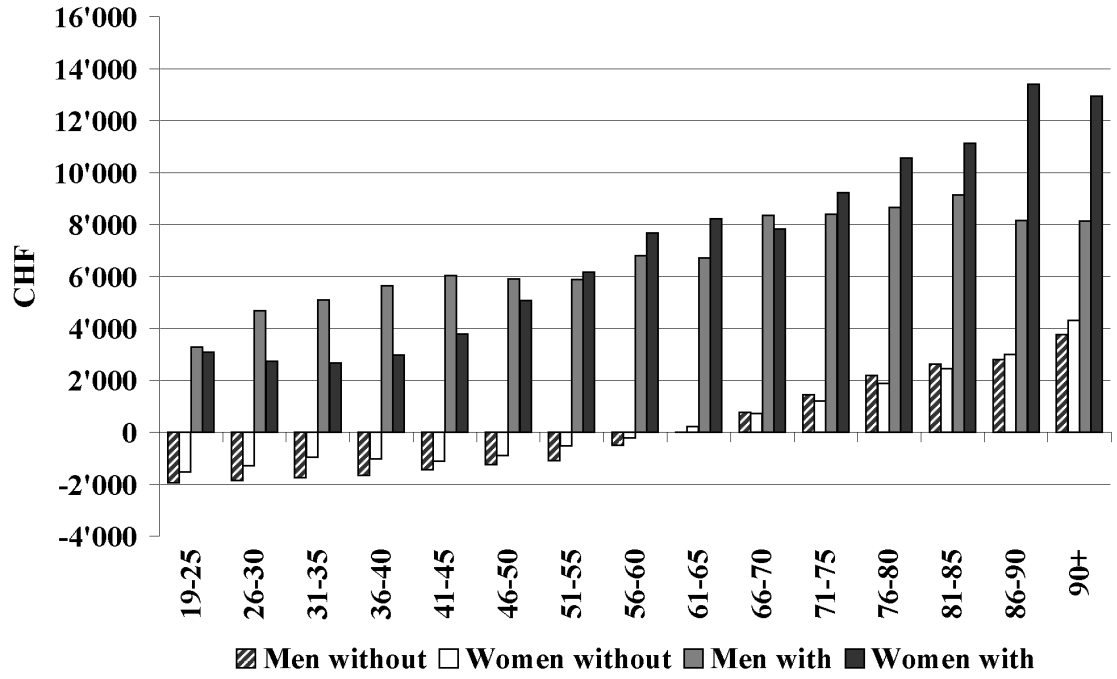


Figure 4.1: Cross-subsidies by age group, persons without / with hospitalization or living in nursing home during the previous year, canton of Zurich, CHF (2005)

burdened to the benefit of those beyond age 60, the new formula makes hospitalization the crucial determinant of CS values. Figure 4.1 illustrates the case of the canton of Zurich. The additional adjuster would cause persons with a hospital stay in 2004 to be cross-subsidized in 2005 regardless of age or gender. It is remarkable that until age 50, men benefit more (due to a higher rate of hospitalization) than do women. With the additional criterion, women contribute less to and receive more from CS, with the age group 51-55 marking the crossover (except for the 66-70 bracket).

4.4 Limiting the Volume of Risk Adjustment

The volumes of CS and RA have increased significantly since 1996 (see Table 4.4), when the new law on health insurance came into effect (and along with it, RA). Both CS growth (80.1 percent up to 2005) and RA growth (126.8 percent) have consistently outstripped the 60.9 percent of net HCE (defined as payments by health insurers with copayment deducted). While HCE growth, ceteris paribus, increases incentives

of both favorable risks and insurers to engage in risk selection, it evidently fails to fully explain the development of CS and RA volumes. Those discrepancies can be interpreted in two ways.

On the one hand, rapid growth of RA volume may be the consequence of increases of risk-selection effort in excess of HCE growth. On the part of insurers, this is the predicted response to increasing discrepancies between risk-based premiums and imposed community-rating (see Zweifel, 2007 for a theoretical development). According to Table 4.2, this may well have been the case since the increase in variance of HCE (reflected by the standard error of cross-subsidies) is mainly associated with age (and hence, higher HCE). With HCE growing rapidly from 1996 to 2005 (see Table 4.4 again), its variance likely grew as well and with it the gap between actuarially fair and community-rated premiums. On the part of consumers, favorable risks may well have stepped up their efforts at avoiding the rapidly increasing burden of CS. One way was to opt for higher deductibles and MC options because insurers, while not

Table 4.4: Volumes of cross-subsidization (CS) and risk adjustment (RA), CHF (1996-2005)

Year	Vol. CS (CHF mn.)	Change (%)	Vol. RA (CHF mn.)	Change (%)	Vol. HCEn* (CHF mn.)	Change (%)
1996	2,700	–	530	–	10,779	–
1997	2,920	+8.2	531	+0.2	11,361	+5.4
1998	3,195	+9.4	609	+14.7	11,927	+5.0
1999	3,366	+5.3	660	+8.4	12,431	+4.2
2000	3,575	+6.2	735	+11.4	13,190	+6.1
2001	3,826	+7.0	853	+16.1	13,986	+6.0
2002	4,009	+4.8	937	+9.8	14,593	+4.3
2003	4,250	+6.0	1,009	+7.7	15,336	+5.1
2004	4,568	+7.5	1,103	+9.3	16,308	+6.3
2005	4,864	+6.4	1,202	+9.0	17,353	+6.4
Avg. change		+6.7		+5.4		
Total change	+80.1		+126.8		+60.9	

Note: HCEn*: HCE - deductibles - copayment, "net HCE"; p.c. = per capita, CHF 1 ≈ USD 0.83 at 2007 exchange rates, Source: Joint Organization KVG (2005), Federal Office of Public Health (2007)

permitted to pass on the full savings, still could pass on more than the "true" savings after deduction of risk-selection effects (which amount to between one- and two-thirds of the full savings in the case of MC, as estimated in Lehmann and Zweifel, 2004). As stated in Section 4.2, both contractual variants gained a great deal of market share just between 2001 and 2005. This interpretation points to activities designed to circumvent premium regulation. They could be reined in by perfecting the RA scheme. Recall that with perfect RA, insurers would not be able to offer a share of their "fake" savings to consumers who seek to dodge the cross-subsidy. This consideration motivated the Swiss parliament to pass a refinement of the RA formula by including the hospitalization criterion.

On the other hand, RA seems to have become a redistributive scheme with a life of its own. Indeed, the CS volume in favor of the old grew even faster than the 80 percent shown in Table 4.4 (not evidenced here), which was not anticipated. A refinement causes CS and RA volumes to increase even more substantially, as evidenced in Section 4.3. The consequence is to increasingly shelter insurers from financial risk undermining their incentive to improve efficiency. These considerations motivated the Swiss Council of States and the Swiss National Council to discuss capping the CS volume at its 2005 value (CHF 4.8 bn.).

As a certainly second-best measure, capping the volume of RA (or indeed CS) is considered below. A simple limit in nominal terms would even have the advantage of increasing insurers' risk exposure over time, forcing them to step up their efficiency-enhancing efforts. It could be imposed at three levels, the aggregate (broken down to insurers according to market share, which may be changing over time), the individual insurer (fixed over time), and the consumer (limiting directly the amount of CS). Only a cap on total CS volume (see second column of Table 4.4) will be considered because it is invariant to changes in market share and structure. A question that naturally arises at this point is who bears the consequences of a cap. Three parties can be identified.

- *The individual insured.* A cap on CS volume causes premiums to converge towards risk-rated values, causing CS values to decrease (also leading to a mitigation of moral-hazard effects (see e.g. Zweifel and Breuer, 2006)).

- *The insurers.* In a system with community-rated premiums, RA is introduced to eliminate (or reduce) incentives for risk selection. Capping its volume (directly or indirectly through limiting CS) causes insurers not to be fully compensated anymore for enrolling unfavorable risks. In the Swiss context, there are two predicted responses. One is to eschew high risks, using known means such as losing application forms. The other is to form conglomerates with a shared sales office (see Van de Ven et al., 2003). Potential clients are quickly assessed on the telephone and assigned to a low-premium affiliate if found a low risk or a high-premium one otherwise. While the lower risks are happy to accept, the high ones often prefer accepting the higher premium to overcoming the hurdles erected by conventional competitors with their lower community-rated premium. Although this practice is not in the spirit of the law, it is legal because each affiliate of the conglomerate does charge a uniform premium.

- *The cantons.* Capping RA increases the financial burden of cantons with an unfavorable risk structure because the cantons pay part of the subsidies to those (in part high-risk) citizens whose premiums exceed a certain share of their income. If the RA volume is capped, high risks become even more unfavorable for health insurers than before. This is especially true for the low-cost insurers of a conglomerate. While the conglomerate cannot legally cancel the policy, it can urge consumers to move to a high-cost affiliate that also charges higher premiums. This causes the share of income devoted to health insurance, and hence premium subsidies to increase, especially if there is a correlation between low income and high risks. However, the federal government is affected as well through matching grants.

4.5 Optimizing the Cap on the Volume of Risk Adjustment

The preceding section has shown that capping the volume of RA has opportunity costs in terms of increased risk-selection efforts on the part of both insurers and consumers. This gives rise to an optimization problem: How is the cap to be allocated to minimize its opportunity cost? The development below focuses on insurers' incentive for risk selection, neglecting changes in consumer behavior in response to reduced cross-subsidies.

Swiss RA is based on age and gender. Its values are calculated in the following way,

$$RA_{a,g} = \bar{L}_{a,g} - \bar{L}, \tag{4.1}$$

where a and g are indexes for age and gender categories, $RA_{a,g}$ is the payment to ($RA_{a,g} < 0$) or from ($RA_{a,g} > 0$) RA in group (a, g), $\bar{L}_{a,g}$ is average HCE in group (a, g), and $\bar{L}$ is average HCE paid by insurers in the population as a whole. The volume of CS (V) can then be calculated as in Equation (4.2),

$$V = \{\sum_{a=1}^{15} \sum_{g=0}^{1} |RA_{a,g}| n_{a,g}\}/2, \tag{4.2}$$

where $n_{a,g}$ is the number of insured of risk group with age a and gender g. RA payments are considered in absolute values to avoid canceling out of positive and negative values. However, this makes the division by two necessary to avoid double counting.

Favorable risks contribute to the insurer's margin, which can be used to cover the deficits generated by unfavorable risks. The insurer is exposed to a higher risk of insolvency if these deficits are large. Reserves can be used to ensure solvency, but too many outliers endanger the economic survival of the insurer. There are several methods for analyzing the importance of such outliers, such as value-at-risk or expected shortfall (see Hull, 2006). However, the easiest way to proceed is to analyze

the variance of HCE falling on the insurer.

If one considers age and gender as the only determinants of HCE (which is in accord with current Swiss RA), then variance in HCE across these groups is given by

$$s_L^2 = \frac{\sum_a \sum_g (\bar{L}_{a,g} - \bar{L})^2 n_{a,g}}{\sum_a \sum_g n_{a,g}} \tag{4.3}$$

where $\bar{L}$ is total average HCE in a specific canton (recall that RA is calculated in each canton separately), and $\bar{L}_{a,g}$ is average HCE of a specific age and gender cell. As of 2005, s_L^2 is estimated for the canton of Zurich at CHF 6.0 mn. RA thus serves to reduce the variance of HCE falling on insurers (and therefore mitigate the incentive to "skim the cream"). This can be seen by plugging Equation (4.1) into Equation (4.3) and rearranging terms,

$$s_L^2 = \frac{\sum_a \sum_g RA_{a,g}^2 n_{a,g}}{\sum_a \sum_g n_{a,g}} \tag{4.4}$$

and hence,

$$(\sum_a \sum_g n_{a,g}) s_L^2 = \sum_a \sum_g RA_{a,g}^2 n_{a,g}. \tag{4.5}$$

Equation (4.5) shows that with a constant number of individuals in each age and gender cell, RA values must be increasing with increasing differences in HCE between groups. This of course serves to increase RA and CS volume as well. If age and gender would be the only determinants of HCE (i.e. if insurers had no private information about individuals, contrary to the analysis by Shen and Ellis, 2002a, Shen and Ellis, 2002b), then risk adjustment would eliminate all risk induced by community-rating. Prior to capping the volume of RA, the variance borne by the health insurer (s_{HI}^2) would be zero, s_{HI}^2=0. This evidently does not hold in the present context because RA in Switzerland is far from perfect (Beck et al., 2006).[2] Whatever the initial value of s_{HI}^2, capping the volume of RA causes it to increase. The objective therefore is to

[2]Note, however, in both cases (s_{HI}^2=0 or s_{HI}^2=c, where c is a constant) the results for the optimization problem is the same, since a constant cancels out when taking derivatives.

minimize this increase, subject to CS volume not exceeding the cap $\bar{V}$. RA payments $\widehat{RA}_{a,g}$ are the decision variables in the problem,

$$\min_{\widehat{RA}_{a,g}} \hat{s}^2_{HI} - s^2_{HI} \qquad \text{s.t.} \qquad V \leq \bar{V}, \tag{4.6}$$

with $\hat{s}^2_{HI}$ denoting the variance when volume is capped. Of course the optimization must take into account that risk adjustment is zero sum. This however is always achieved since positive and negative RA values cancel out.

If volume is defined as in Equation (4.2), optimization is difficult due to absolute values. An alternative approach is therefore taken here. First, the positive half variance $\hat{s}^2_{HI+}$ (with the restriction on CS volume), and then, the negative half variance $\hat{s}^2_{HI-}$ (with the same restriction on CS volume) is minimized, ensuring that RA values sum to zero,

$$\min_{\widehat{RA}_{a,g}} [(\hat{s}^2_{HI+}) + (\hat{s}^2_{HI-})] - s^2_{HI}. \tag{4.7}$$

Because s^2_{HI} is predetermined, it is obvious that only the terms in brackets are relevant. The first term can be broken down as shown in Equation (4.8),

$$\min \hat{s}^2_{HI+} = \frac{\sum_a \sum_g (x_{a,g} - \bar{x})^2 n_{a,g}}{\sum_a \sum_g n_{a,g}} \qquad \text{s.t.} \qquad x_{a,g} > \bar{x}. \tag{4.8}$$

The symbols are defined as follows,

$$\begin{aligned} x_{a,g} &= (L_{a,g} - \widehat{RA}_{a,g}) \\ \bar{x} &= \bar{L} \\ x_{a,g} - \bar{x} &= (L_{a,g} - \widehat{RA}_{a,g} - \bar{L}) \\ &= (RA_{a,g} - \widehat{RA}_{a,g}). \end{aligned} \tag{4.9}$$

Here, $\widehat{RA}_{a,g} > 0$ if the insurer receives a payment from the RA scheme and consumers in the (a, g) cell receive a cross-subsidy. Conversely, $\widehat{RA}_{a,g} < 0$ if it pays into the

scheme (and low-risk consumers bear a cross-subsidy).

Since the restriction $V \leq \bar{V}$ always holds as an equality in Equation (4.6), the problem can be solved using a Lagrangian,

$$\begin{aligned} \min_{\widehat{RA}_{a,g}} Z &= (\hat{s}_+^2) - \lambda(\sum_a \sum_g \widehat{RA}_{a,g} n_{a,g} - \bar{V}), \qquad \forall \quad \widehat{RA}_{a,g} > 0 \qquad (4.10) \\ &= \frac{\sum_a \sum_g (RA_{a,g} - \widehat{RA}_{a,g})^2 n_{a,g}}{\sum_{a=1}^{15} \sum_{g=0}^{1} n_{a,g}} - \lambda(\sum_a \sum_g \widehat{RA}_{a,g} n_{a,g} - \bar{V}), \end{aligned}$$

where the subscript HI is dropped for simplicity. The solution to this problem shows how positive payments received from the RA scheme are optimally reduced. Payments into the scheme ($\widehat{RA}_{a,g} < 0$) are fully analogous. The first-order conditions are,

$$\begin{aligned} \frac{\partial Z}{\partial \widehat{RA}_{a,g}} &= \frac{-2(RA_{a,g} - \widehat{RA}_{a,g}) n_{a,g}}{\sum_a \sum_g n_{a,g}} - \lambda n_{a,g} = 0 \qquad (4.11) \\ \frac{\partial Z}{\partial \lambda} &= \sum_{a=1}^{15} \sum_{g=0}^{1} \widehat{RA}_{a,g} n_{a,g} - \bar{V} = 0. \end{aligned}$$

This is a system of linear equations in $\widehat{RA}_{a,g}$ and λ that has full rank and can therefore be solved. An example with four risk classes is given as follows.

Assume a hypothetical RA scheme distinguishing four groups $i = 0, 1, 2, 3$ with n_i the number of individuals in that group and n the overall number of individuals. Let two groups (0 and 1) have below-average and two (2 and 3), above-average expected HCE. RA_i indicates RA payments for each group. The first-order conditions for negative payments (i.e. payments to the RA scheme, groups 0 and 1) are,

$$\begin{aligned} \frac{\partial Z}{\partial \widehat{RA}_0} &= \frac{-2(RA_0 - \widehat{RA}_0) n_0}{n} - \lambda n_0 = 0 \qquad (4.12) \\ \frac{\partial Z}{\partial \widehat{RA}_1} &= \frac{-2(RA_1 - \widehat{RA}_1) n_1}{n} - \lambda n_1 = 0 \\ \frac{\partial Z}{\partial \lambda} &= \widehat{RA}_0 n_0 + \widehat{RA}_1 n_1 - \bar{V} = 0. \end{aligned}$$

Now λ can be solved for from the first FOC,

$$\lambda = \frac{-2(RA_0 - \widehat{RA_0})}{n}. \tag{4.13}$$

Equation (4.13) shows the determinants of the opportunity cost caused by the cap. First, the greater the population at risk (n), the smaller this cost. Second, the greater the difference between RA payments with and without the cap ($RA_0 - \widehat{RA_0}$), the higher this cost. In addition, the system (4.12) can be solved to yield,

$$\begin{aligned} (RA_0 - \widehat{RA_0}) &= (RA_1 - \widehat{RA_1}) \\ \widehat{RA_0} n_0 + \widehat{RA_1} n_1 &= \bar{V}. \end{aligned} \tag{4.14}$$

It is evident that the optimal reductions of RA values are the same across risk categories. Solving this system of two equations in the two unknowns yields the following solution payments to the RA scheme,

$$\begin{aligned} \widehat{RA_0} &= \frac{\bar{V} - (RA_1 + RA_0)n_1}{n_0 + n_1} = \frac{\bar{V}h_1}{n_1} - (RA_1 - RA_0)h_1 \\ \widehat{RA_1} &= \frac{\bar{V} - (RA_0 + RA_1)n_0}{n_0 + n_1} = \frac{\bar{V}h_0}{n_0} - (RA_0 - RA_1)h_0 \end{aligned} \tag{4.15}$$

with h_i noting the share of group i in the subpopulation with below-average HCE. Therefore, the optimal new RA values are

- the lower, the lower the cap is set;
- the lower, the greater the positive difference in RA values prior to the limitation (e.g. $RA_1 > RA_0$);
- the higher, the greater the negative difference in RA values prior to the limitation (e.g. $RA_0 < RA_1$);
- the higher, the higher the group's population share h_i (even for small n_0 since $\bar{V} >> (RA_1 - RA_0)$).

The payments received from the RA scheme can be derived in an analogous way.

4.6 Consequences of Capping Risk Adjustment

As argued in Sections 4.1 and 4.2, risk-selection behavior is ultimately driven by the amount of CS contained in contributions to health insurance. And in the case of Switzerland, the political debate revolving around RA has focused on the CS rather than the RA volume. For these reasons, this section cites more CS rather than RA figures.

4.6.1 Theoretical Considerations

The question as to the optimal value of the cap cannot be addressed in this paper. It requires knowledge of citizens' willingness-to-pay for avoiding risk-selection efforts by health insurers while keeping community-rated premiums. Experimental evidence concerning willingness-to-pay for attributes of health services provision has been presented (in e.g. Telser et al., 2004, Zweifel et al., 2006, and concerning attributes of health insurance, in Becker, 2006 and Becker et al., 2007). However, willingness-to-pay for maintaining community-rated premiums has not been measured to the knowledge of the authors. As a second-best solution, parliament could decide on the value of the cap, assuming that politicians represent the preferences of the population.

While the political debate has focused on the national level, cantons will likely be affected as well. As evident from Equations (4.9) and (4.11), the opportunity cost of a cap on RA is linked to the dispersion of HCE, which varies between cantons. If CS and hence RA volumes were to be limited, many citizens with low incomes would have to pay higher premiums. This creates political pressure for increased redistribution through premium subsidies. More generous cantons would be more prone to increasing their subsidies, which are matched by the federal level, where a substantial amount of redistribution between cantons takes place. Therefore, a limit on CS and RA volumes is likely to induce a certain amount of CS between cantons.

Focusing on the opportunity cost of a cap in terms of incentives when an additional adjuster is introduced, there are two effects to be distinguished. First, since HCE is now

predicted more precisely, the variance of HCE borne by insurers decreases and with it risk-selection effort. Second, unless the additional adjuster exhibits perfectly negative correlation with the existing ones, CS and RA volumes must increase. However, the incidence of hospitalization increases with age and is higher among women than men; therefore volumes increase. This increase affects the opportunity cost, depending on the situation.

- The benchmark case is no cap, combined with the introduction of the hospitalization criterion into the RA formula. This simply reduces the variance in HCE to be borne by insurers, thus mitigating incentives for risk selection (see Equation (4.4)).

- The cap is imposed but not binding initially; it becomes binding with the introduction of the additional RA criterion. Therefore the opportunity cost of the cap was zero at the beginning. It would become positive but still small if the CS volume had to be reduced from CHF 4.6 to 4.5 bn. since the effect on insurer behavior is still limited. However, a future reduction from CHF 6 to 4.5 bn. (say), would cause opportunity cost to rise (see Equation (4.13)).

- The third alternative is the introduction of the additional RA criterion when the cap is already binding. On the one hand, this would serve to reduce the volatility of HCE falling on health insurers. On the other hand, the restriction on CS volume becomes even more binding. The first effect mitigates incentives for risk selection, while the second strengthens them. The net effect remains ambiguous (see Equations (4.4) and (4.13), respectively).

Note that CS and RA volumes can always be reduced by permitting health insurers to charge premiums that are more in line with true risk. For example, suppose smokers pay an additional premium of CHF 50 per year. This would decrease the difference between HCE and premium revenue by CHF 50 ceteris paribus and hence the variance of payments and with it the risk to be borne by the health insurer. Incentives for risk selection decrease. The advantage of more risk-rated premiums is that RA volume declines endogenously without inducing more efforts at risk selection; its drawback is

deviating from community-rating. Conversely, the advantage of capping CS and RA volumes is that community-rating can be retained, while its downside is that incentives for risk selection are strengthened.

4.6.2 Empirical Illustration

Since the unit of reference for Swiss RA is the canton, the effects of limiting its volume can be exemplified by using data for the canton of Zurich, assuming a decrease of CS volume from the estimated nationwide CHF 5.375 bn. level to CHF 4.5 bn. The estimate is derived from pitting expected HCE at the individual level against the mean applicable to each of the 25 health insurers operating in the Canton of Zurich. Expected HCE was estimated using a two-part model along the lines of Steinmann et al. (2007), pooling the data provided by the three health insurers. However, dummy variables reflecting RA cells replaced continuously measured age. Other dummies are

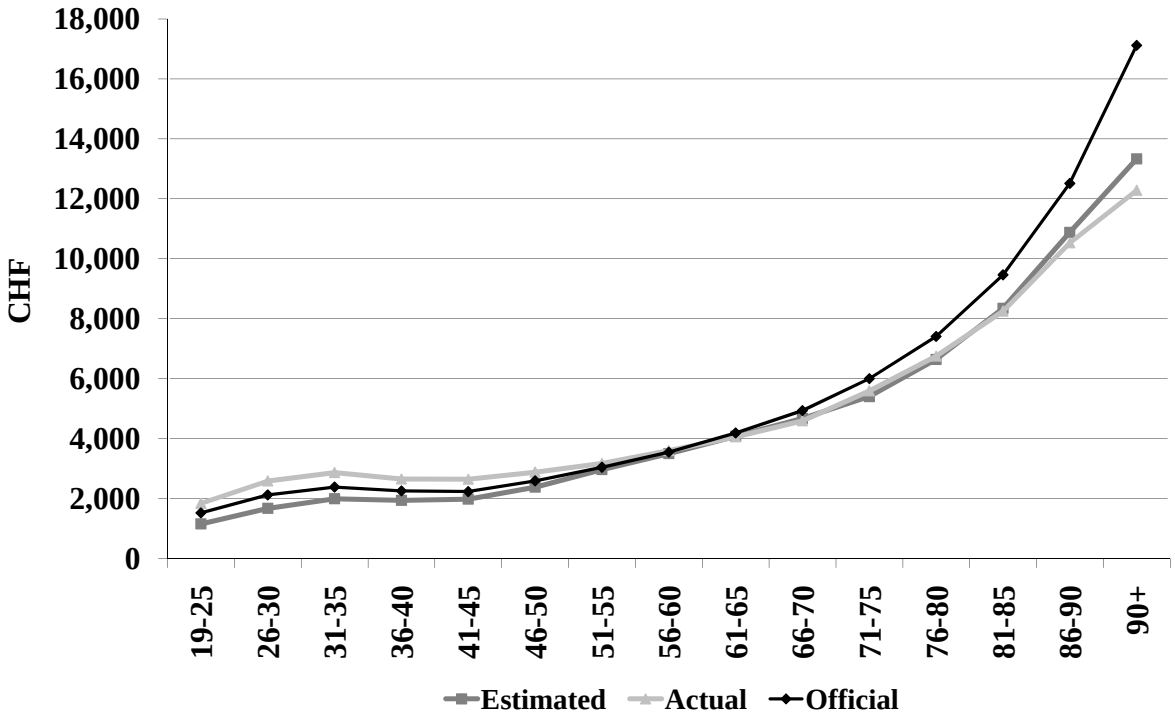


Figure 4.2: HCE of women by age, actual, estimated (three insurers), and official, Switzerland, CHF (2005)

female gender and canton of residence.

Figure 4.2 shows that the estimated age profile of HCE closely reflects that of the three insurers ("Actual") that provided the data. However, the increase of HCE with higher age at the national level ("Official") is still underestimated. Overall, estimated HCE seem to fit the Swiss average sufficiently well to derive estimates of the effect a cap on CS volume may have, assuming that it is imposed in a way as to minimize its opportunity cost (as expounded in Section 4.5).

Table 4.5 illustrates the effects of a reduction of CS from CHF 5.375 bn. to CHF 4.5 bn. for the canton of Zurich (number of insured: 1,032,600). Initially the youngest age class of women pays a premium that exceeds expected HCE by some CHF 1,600, used to finance the higher HCE of Zurich residents above 50 years of age. Capping the CS volume would reduce this excess by CHF 236. On the other hand, the highest age group of females currently receives more than CHF 11,000 as a

Table 4.5: Capping the volume of cross-subsidies, canton of Zurich, CHF (2005)

	Women			Men		
Age group	Without cap	Change	With cap	Without cap	Change	With cap
19-25	-1,618	+236	-1,382	-2,347	+236	-2,111
26-30	-1,103	+236	-867	-2,290	+236	-2,054
31-35	-772	+236	-536	-2,191	+236	-1,955
36-40	-830	+236	-594	-2,073	+236	-1,837
41-45	-789	+236	-553	-1,800	+236	-1,564
46-50	-350	+236	-114	-1,479	+236	-1,243
51-55	283	-283	0	-548	+236	-312
56-60	855	-435	421	223	-223	0
61-65	1479	-435	1,045	1,223	-435	788
66-70	2,126	-435	1,691	2,085	-435	1,650
71-75	2,886	-435	2,451	3,301	-435	2,866
76-80	4,191	-435	3,757	3,276	-435	2,841
81-85	5,983	-435	5,549	4,851	-435	4,417
86-90	8,643	-435	8,209	7,162	-435	6,728
90+	11,250	-435	10,816	9,419	-435	8,985

Note: CHF 1 ≈ USD 0.83 at 2007 exchange rates

cross-subsidy. This would be reduced by CHF 435. An exception is the class of 51-55 year old women, which changes from receivers to neutral.[3] These amounts are to be compared with the average Zurich premium, which was about CHF 4,000 in 2005.

Naturally, capping the CS volume has an impact on insurers operating in the canton of Zurich as well. First, the decrease in CS values evidenced in Table 4.5 implies an increase in the deviations between actual and average HCE falling on them. Therefore, the variance of HCE borne by them (s^2_{HI} in Section 4.5) is bound to increase, very likely triggering additional risk-selection efforts on their part. This effect would even be more pronounced if the cap on CS and RA volumes were to be imposed in a non-optimal way. Second, the amount of RA transferred between insurers would fall. This effect was estimated in the following way.

Equation (4.1) for determining RA values was implemented using the estimated HCE function to assign HCE values to age/gender/hospitalization cells of all 25 insurers operating in the Canton of Zurich. Next, the pool over which RA is defined was restricted to these 25 insurers. For the canton of Zurich, the error incurred is small because out of the 1,032,600 insured, only 172,671 do not belong to one of three insurers considered or one of the additional 22 sampled. Moreover, the resulting underestimation should not influence the percentage reduction much since it affects both uncapped and capped values. With the 25 insurers having a nationwide market share of 60 percent, the resulting total of RA values is scaled up accordingly to obtain CHF 1.123 bn. as the national estimate prior to imposing a cap (see Table 4.6).

For simplicity, the simulated RA values with the CS cap were not optimized (implementing the Equation system (4.11) would constitute a research paper of its own) but simply allocated evenly to the age/gender/hospitalization status cells of insurers according to their shares in current RA volume. Thus, the cap on CS volume would

[3] According to Equation (4.14), optimal reduction of RA values is the same across risk categories. However, since the class of 51-55 year old women and 56-60 year old men changes from receiver to payer, RA values have to be adjusted correspondingly.

Table 4.6: Change of CS and RA volumes, CHF bn. (2005)

	Without cap	With cap	Change
CS volume between consumers	5.375	4.500	-16.3%
RA volume between insurers	1.123	0.982	-12.5%

Note: CHF 1 ≈ USD 0.83 at 2007 exchange rates

reduce the RA volume by an estimated 12.5 percent. This is markedly less than the 16.3 percent reduction of CS volume because a great deal of CS occurs between consumers enrolled in a given fund.

4.7 Conclusions

This contribution addresses the conflict of interest arising in the context of imperfect risk adjustment (RA). On the one hand, a refinement of the RA formula would weaken health insurers' incentive to engage in risk selection (given that they are subject to community-rating). On the other hand, unless fully prospective, RA undermines their quest for efficiency. There are three novel aspects to this paper. First, it adopts standard economic theorizing by distinguishing between the insurers who pay (through their contributions to the RA scheme) and the favorable risks who ultimately bear these cross-subsidies (CS), amounting to the difference between their actual and their actuarially fair premium. Second, as a rough-and-ready measure to expose insurers to a degree of financial risk, a cap on RA is considered and the resulting optimization problem studied. The issue is to structure the reduced RA values in a way as to minimize the increase of HCE variance borne by the insurer (and hence risk-selection effort). Third, the study simulates the consequences of a cap (on total CS rather than RA volume) both for consumers and insurers.

In a first step, data provided by three Swiss health insurers is compared to official nationwide averages to assess their representativeness. Overall, the data seems to accord with official statistics to a sufficient degree to justify more detailed investigation.

Next, the refinement of the Swiss RA formula effective 2012 is considered. The inclusion of the additional criterion, "Hospitalization or living in a nursing home during the previous year", is found to inflate CS volume by 40 percent on average and to cause age and gender to lose importance as risk adjusters throughout. This increase of CS (and hence RA) volume contradicts the objective of enhancing insurers' incentives for efficiency by exposing them to more financial risk, to be achieved by a cap on RA volume.

This conflict of interest gives rise to the optimization problem, "Minimize the HCE variance falling on the insurer (and therefore the incentive for risk selection), subject to RA volume not exceeding a politically determined level". The optimal solution calls for a uniform reduction of positive and negative RA values the amount of which depends on existing differences between groups in terms of RA values and their population shares.

A simulation extrapolating from one Swiss canton shows that a reduction of CS volume to CHF 4.5 bn. (by 16 percent) at the national level would reduce the RA volume between insurers by an estimated 12.5 percent. The optimized CS burden would drop slightly for those up to age 55, juxtaposed by a reduction of CS in favor of those above 55. However, HCE variance falling on insurers would increase, strengthening their incentives for risk selection.

This research is subject to several limitations. First, the refinement of the RA formula considered is one among many, e.g. the inclusion of diagnostic information. Second, capping CS (or RA) values to push insurers towards efficiency is certainly second best. This objective could be achieved at a lower opportunity cost if alternatives such as optimized cut-off points in the HCE distribution (beyond which RA sets in) were considered, with potentially quite different implications for CS values between consumers. Third, behavior of insurers was assumed to be driven by the HCE variance falling on them, while that of consumers, by the gap between the actual and the actuarially fair premium. Especially the latter assumption can be criticized for its neglect of fairness

considerations. Still, this research does address some of the issues raised by a rough-and-ready measure that may appeal to politicians, such as simply capping the amount of RA (or CS) volume.

Acknowledgments

The authors gratefully acknowledge helpful suggestions by Konstantin Beck (University of Zurich and CSS Insurance, Lucerne, Switzerland), Willard G. Manning (University of Chicago, Illinois, USA), the Health Economics Working Party of the Verein für Socialpolitik at the conference 2007, and by two anonymous referees. Furthermore, special thanks go to the three health insurers that provided data for this study.

Chapter V
Fine-Tuning of Health Insurance Regulation: Unhealthy Consequences for an Individual Insurer

Johannes Schoder, Michèle Sennhauser
and Peter Zweifel

Published in

International Journal of
the Economics of Business,
Vol. 17 (3), p. 313-327, November 2010

5 Fine-Tuning of Health Insurance Regulation: Unhealthy Consequences for an Individual Insurer

5.1 Introduction

When premiums are mandated to be independent of risk, competitive health insurers have an incentive to select clients whose future expected health care expenditure (HCE) does not exceed their contribution. This consideration has induced secondary regulation in the guise of risk adjustment (RA) schemes. Basically, RA makes insurers with an above-average share of favorable risks pay into a fund, whose proceeds are used to cross-subsidize those insurers with many unfavorable risks. The design of an optimal RA formula is a widely discussed topic (see for example Lamers, 1999, Ellis and Van de Ven, 2000, Glazer and McGuire, 2002, Lamers and Van Vliet, 2003a, Lamers and Van Vliet, 2003b, Van de Ven et al., 2004, Beck et al., 2006, Jack, 2006, Zweifel and Breuer, 2006, and Van de Ven et al., 2007). The RA formulas for Medicare in the United States and the Netherlands are being refined continuously (see e.g Douven, 2007 and Calfo, 2009). However, so far the consequences of this fine-tuning of regulation for the risk management (RM) of insurers seem to have been neglected.

This contribution contains a case study from Switzerland, a country that relies on competitive health insurance in a way similar to the US and the Netherlands. A RA scheme was introduced in 1993, using the two criteria age and gender only. Effective 2012, the RA formula will include a third indicator of high risk, viz. "Hospitalization of more than three days or living in a nursing home during the previous year" (see Spycher, 2000). While this choice is largely dictated by service providers' refusal to

pass on diagnostic information to health insurers, it does have several recommendable features in that it (1) has significant predictive power (see Holly et al., 2003 and Beck, 2004), (2) relates to a previous period so does not undermine insurers' effort at controlling health care cost, and (3) can be measured at little administrative expense.

Refinement of the RA formula has gone much further in other countries. In the United States, the CMS hierarchical condition categories model (CMS-HCC) has been in use with Medicare since 2004. It uses diagnoses from all clinical encounters, regardless of whether they are inpatient or outpatient (see Pope et al., 2004). In the Netherlands diagnostic cost groups (DCGs) and pharmacy-based cost groups (PCGs) are used as high-risk indicators.[1] These reforms have their costs and benefits. On the benefit side, risk-selection efforts by health insurers are reduced if the net cost of medical care falling on them is increasingly equalized across risk types. Moreover, this net cost does not depend anymore on whether the insured were hospitalized or not. On the cost side, these refinements of RA not only require more accounting effort on the part of both insurers and providers but also increase proneness to error[2]. Moreover, they create incentives for up-coding diagnoses (for an explicit analysis of advantages and disadvantages in the case of United States Medicare, see Pope et al., 2000 and Kominski, 2007).

The purpose of this paper is to point out another cost of RA refinement. Indeed, it may boost payments into the RA scheme to an extent as to jeopardize the economic survival of an otherwise viable health insurer, posing a great challenge to its RM. Now insolvency and hence market exit of an insurer who only survived thanks to cream skimming may be considered to be efficiency enhancing. However, this case study deals with an innovative health insurer, who had successfully implemented Managed Care (MC) to lower rates of hospitalization. Bankruptcy of such an insurer would

[1] They are derived from the diagnoses related to prior hospitalization and prior use of prescription drugs, see e.g. Van de Ven and Schut (2008).

[2] In the Netherlands, the complexity of processing the data and money flows led to errors in the calculation of the ex-ante risk-adjusted capitation payments, resulting in a loss of Euro 247 million (mn.), falling on taxpayers (see Douven, 2007).

have to be considered inefficient.

The evidence comes from simulating payments for a particular health insurer A into the RA scheme applying the old and the new formula. These simulations predict that A's payments would have increased significantly, attaining between 9 and 13 percent of premium income. Extra payments of these magnitudes would have seriously endangered insurer A's economic survival, leading to a cumulative loss in excess of CHF 250 mn. (CHF 1 $\approx$ USD 1 at 2010 exchange rates) over three years. While A's RM response cannot be predicted, there are two main alternatives. One is to enlist unfavorable risks, as intended by the regulator. The other is to extend hospital stays from three to four days. This strategy would have decreased this insurer's RA payments by an estimated 11 percent in 2007. The consequences would be unhealthy for taxpayers (who subsidize hospital cost), employers (who lose workdays), and patients (who lose quality of life). While not directly transferable to other countries with competitive health insurance (such as the United States, but also Germany, Israel, and the Netherlands), the findings of this contribution convey a clear message. Seemingly minor fine-tuning of health insurance regulation has the potential of challenging an insurer's RM, with undesirable consequences for the society.

The remainder of this chapter is structured as follows. Section 5.2 describes the method for calculating risk adjustment values in general and the data basis. In the first part of Section 5.3, RA values are simulated according to the new formula and applied to insurer A. The second part of Section 5.3 analyzes the impact of this regulatory change on insurer A's RM. The chapter concludes with lessons learned from this case study and its implications.

5.2 Simulation of Risk Adjustment Values and Data Basis

5.2.1 Methodology

Traditionally, analysis of RM focuses on payments between health insurers. However, this neglects the fact that payments into the RA scheme are ultimately borne by low-risk consumers while payments from the scheme benefit high-risk consumers. Economic theory has always distinguished between payers and bearers of a cost or levy, in particular in the context of an indirect tax. To see the analogy, consider current Swiss RA with two criteria age and gender only. Define $\bar{P}$ as the community-rated premium, $\bar{L}_{a,g}$, as the average HCE in one of the age-gender cells (a, g) of RA (neglecting administrative expense for simplicity), and $RA_{a,g}$ as the payment to or from the RA scheme. The premium paid by a specific individual i who is a low risk compared to the cohort in the age-gender cell (a, g), and whose expected cost $\mathrm{E}(L_i)$ is thus below average for the specific cell can then be expressed as

$$\begin{aligned} \bar{P} &= \bar{L}_{a,g} + RA_{a,g}, \qquad \text{with } RA_{a,g} > 0 && (5.1) \\ &= \mathrm{E}(L_i) + (\bar{L}_{a,g} - \mathrm{E}(L_i)) + (\bar{P} - \bar{L}_{a,g}). && (5.2) \end{aligned}$$

This particular low risk bears, on top of his or her actuarially fair premium $\mathrm{E}(L_i)$, a cross-subsidy in favor of high risks consisting of two components. The first component is the difference between average HCE of group (a, g) and the individual's expected HCE denoted by $\mathrm{E}(L_i)$; the second, the contribution to the RA scheme $(\bar{P} - \bar{L}_{a,g})$, to be paid by the insurer. The sum of the two will be referred to as cross-subsidization values. As to the second component, the current Swiss RA formula comprises 15 age classes, starting from age 19 to 25 and continuing in 5-year steps. Thus, there are overall 30 RA categories. Since by law risk adjustment must not lead to a cross-subsidization between the 26 cantons (i.e. member states of Switzerland), the RA values are calculated yearly for each canton by the Joint Organization KVG based on data of all Swiss health

insurers (see Joint Organization KVG, 2008). Adopting the insurer's point of view rather than the consumer's now, the RA values are equal to

$$RA_{a,g} = \bar{L}_{a,g} - \bar{L} \tag{5.3}$$

with $\bar{L}$ (= $\bar{P}$ in Equation (5.1) since administrative expense is neglected) denoting average HCE in the canton's population as a whole (see Beck et al., 2006, Chapter 4). Including the criterion "hospitalization"[3] changes Equation (5.3) to

$$RA_{a,g,h} = \bar{L}_{a,g,h} - \bar{L}. \tag{5.4}$$

The subscript h is equal to 1 if a hospital stay in the previous year exceeds three days and 0 otherwise. Average HCE of the respective RA cell, $\bar{L}_{a,g,h}$, now has to be calculated for 60 instead of 30 groups, while $\bar{L}$ remains the same.

The insurer has to contribute to the RA fund for favorable risks ($\bar{L}_{a,g,h}<\bar{L}$). The RA fund uses the proceeds to cover the deficits generated by unfavorable risks ($\bar{L}_{a,g,h}>\bar{L}$). An insurer's total payment (V) into/from the RA fund depends on the composition of its insured over all 26 cantons (c),

$$V = \sum_{c=1}^{26} \sum_{h=0}^{1} \sum_{g=0}^{1} \sum_{a=1}^{15} RA_{a,g,h,c} \cdot n_{a,g,h,c}. \tag{5.5}$$

An insurer receives payments if $V>0$ and contributes to risk adjustment if $V<0$.

5.2.2 Data Basis

For calculating the $RA_{a,g,h,c}$ values in Equation (5.5) for a given health insurer, the cell-specific averages $\bar{L}_{a,g,h,c}$ must be known. Since $RA_{a,g,h,c}$ is not published by the Joint Organization KVG, two different sources are used to analyze the

[3] This is shorthand for "Hospitalization or living in a nursing home during the previous year of four days and more".

impact of the new RA formula on an individual health insurer. The first is constructed by merging individual HCE data provided by three large health insurers in order to calculate the average $RA_{a,g,h,c}$. Ideally it should be representative of all Swiss health insurers. The second data base comes from the one individual Swiss health insurer "A". Both are limited to individuals having mandatory health insurance.

Descriptive Statistics

Data of the three large Swiss health insurers (out of a total of 70 serving a population of 7.5 mn.) is available for the period 2001 to 2005. The sample is well balanced with respect to gender (49.5 percent of women), and average age of adult enrollees (47.4 years in 2005, compared to 47.8 years of the adult population). The market share covered is stable across age classes, amounting to 25 percent on average. With regard to choice of contract, there is a clear trend towards higher deductibles. The three highest deductibles (CHF 1,500, 2,000, and 2,500; CHF 1 $\approx$ USD 1 at 2010 exchange rates) increased in importance from 12 to over 22 percent from 2001 to 2005, which is compared to the official figures of 13 and 23 percent very representative (santésuisse, 2010a). There is a similar trend in favor of MC contracts, reaching a share of 11 percent in 2005 (compared to the Swiss average of less than 10 percent in 2005, see Chapter 4).

The second data source, obtained from A, covers the period 2001 to 2007. With 51.3 percent of women, the sample is almost balanced. A is one of the medium-sized health insurers in Switzerland with a market share of almost 5 percent in 2005. With 47.7 years, average age of A's adult enrollees is slightly higher than the 47.4 years of the three insurers. The clientele of A also tends towards higher deductibles. The share of the three highest deductibles (they are CHF 1,000, CHF 1,500, and CHF 2,500) exceeds the nation-wide average of 22 percent in 2005. MC contracts account for almost 35 percent (2007), double the nationwide average of 16.9 percent (santésuisse, 2010b). This most likely explains A's comparatively low rate of hospitalization (see

Figure 5.3 below).[4] On the whole, A looks like an innovative insurer that encourages MC options, in conformity with stated objectives of Swiss policy makers.

Checking Simulated RA Payments

First, the data provided by the three large health insurers had to be checked for representativeness using the current RA formula. The values for $RA_{a,g}$ were calculated for all 30 cells along with their standard errors according to the methodology described in Section 5.2.1 and compared with the official nationwide values. The insurers on average pay for women aged 19 to 25 more than CHF 1,700 per year (see Figure 5.1 for the canton of Zurich, the leading canton of Switzerland both in terms of GDP and population, and Table 4.2 in Chapter 4, page 78, for all cantons). Conversely, they receive payment for over 90 year old women to the tune of some CHF 8,600. While the fit is good in general, RA contributions by the three insurers are lower than the official figures from age 61 on.

Based on the evidence, one can conclude that the three major health insurers sampled are sufficiently representative of the Swiss population to enable a simulation of the new RA formula based on their data. This conclusion is also supported by the fact that one of the three is a net recipient of payments from the RA scheme, one breaks even, and one is a net contributor to the scheme. Also note that according to Table 4.2 in Chapter 4 (page 78), the standard error and hence variance of RA payments increases with age, reflecting the fact that variance of HCE increases as well. This means that for a risk-averse health insurer, risk-selection effort has a high payoff if focused on older clients. By the same token, however, an insurer like A who counts on having to pay into the RA scheme permanently faces a liability characterized by great risk as its insured population ages.

[4] In the US, MC plans have achieved most cost savings by reducing inpatient hospital use (see Miller and Luft, 1997 and Bindman et al., 2005. For MC cost savings in Switzerland see Lehmann and Zweifel, 2004).

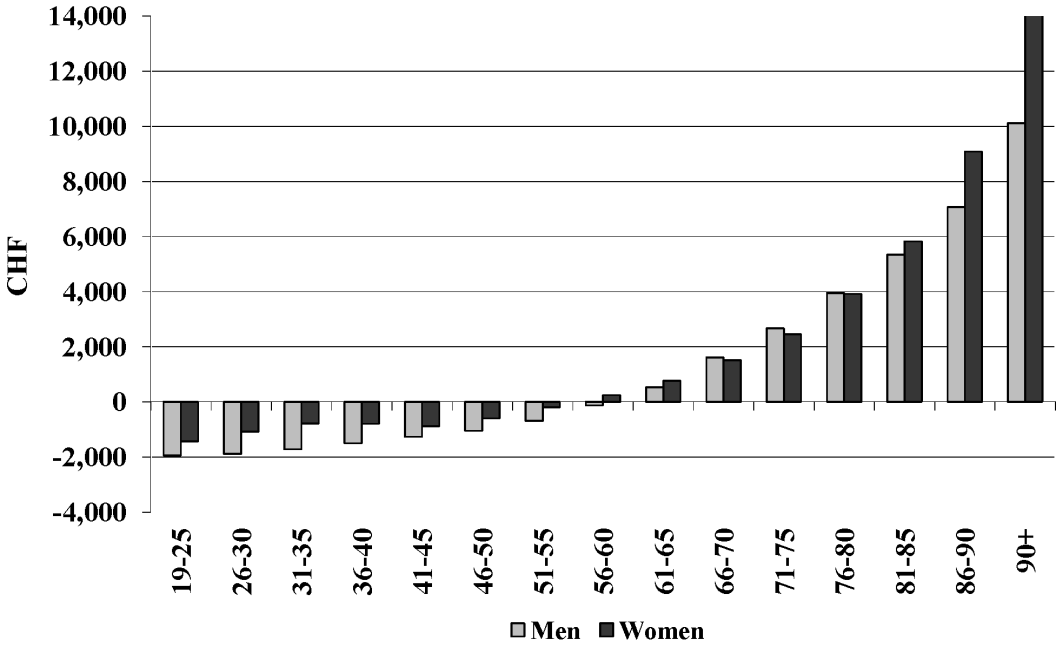


Figure 5.1: Official RA values according to age and gender, canton of Zurich, CHF (2005)

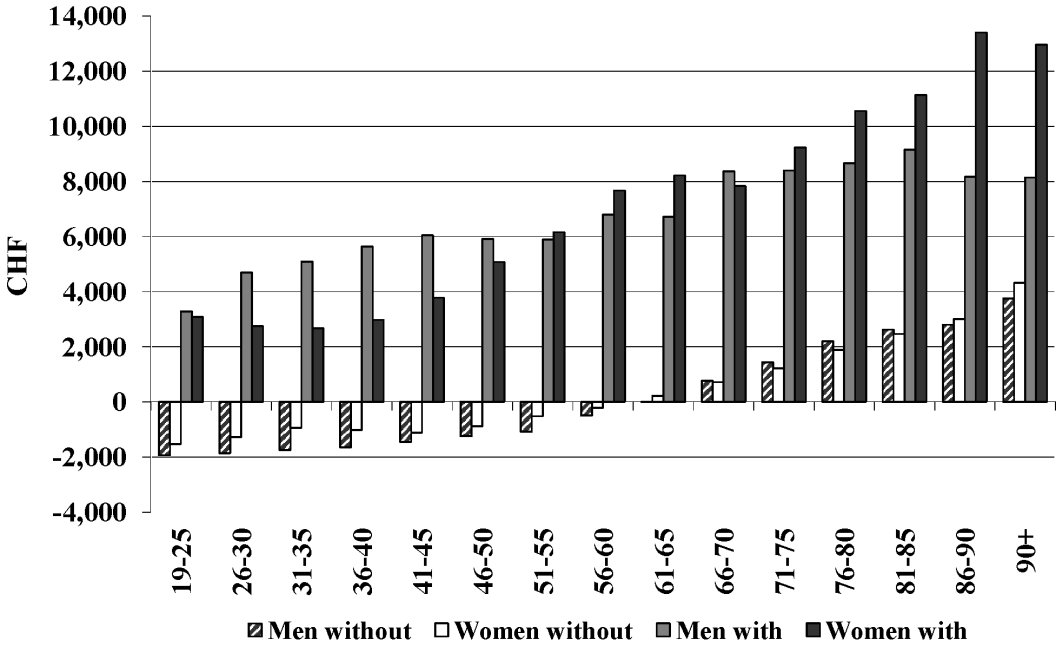


Figure 5.2: Estimated RA values with and without hospitalization according to age and gender, canton of Zurich, CHF (2005)

5.3 Simulating the Impacts of the New RA Formula

In this section, estimated RA values with the new RA formula including hospitalization during the previous year are presented first. Then, the impacts of the regulatory fine-tuning on health insurer A in terms of financial burden and choice of strategy are shown.

5.3.1 Risk Adjustment with the New Criterion

Official RA values grouped according to the additional criterion, "Hospitalization during the previous year" are not available.[5] They have been simulated using the individual HCE data provided by the three major health insurers (see Section 5.2.2). Figure 5.2 illustrates estimated $RA_{a,g,h,c}$ values for the canton of Zurich.

Comparing Figures 5.1 and 5.2 the new formula is seen to induce radical changes. First of all, it causes the amount of cross-subsidization between those without a hospital stay in the previous year to shrink considerably beyond age 70. Conversely, it causes persons with a hospital stay to be cross-subsidized regardless of age or gender. Second, and related to this, the usual age profile ceases to exist. For instance, hospitalized women in the 19 to 25 age group benefit more than the three next older groups, and at the high end, it is the aged 86 to 90 rather than the oldest that benefit most. Among men, the age profile becomes almost level beyond age 70. Third, the per capita amounts now are higher, pointing to a substantial increase in the volume of cross-subsidization. In Chapter 4 the effects of introducing the third criterion on the total volume of cross-subsidization is simulated for 2005. An increase of 40 percent, from CHF 4.13 billion (bn.) to CHF 5.82 bn., or some 12 percent of Swiss HCE is found. Whether this is excessive or not is an issue that cannot be addressed in this paper. However, a change of this magnitude is likely to present a challenge to the RM of at least some health insurers. Whether this is the case of insurer A is the topic of the two subsections below.

[5] Official statistics do show RA values as "RA payments between consumers", but only according to the current RA formula (see Joint Organization KVG, 2008).

5.3.2 Impacts on Risk Adjustment Payments by Health Insurer A

The consequences of adding the new risk adjuster "hospitalization" for health insurer A can be simulated as follows. The volume of payments is calculated as the number of A's customers in a RA cell[6], times the estimated RA value pertaining to that RA cell, and adding up (see Equation (5.5), Section 5.2.1). These calculations are performed using the old and the new RA formula for the years 2005 to 2007. They allow to "postdict" the consequences the new RA formula would have had if already in effect. The results are striking.

- Total payments of A into the RA scheme increase substantially. Under the old formula, they amount to CHF 24.2 mn. in 2005, corresponding to 3 percent of premium income. Had the new RA formula already been in effect, they would have reached CHF 101.6 mn., amounting to no less than 13 percent of premium income. Considering that A operated at a loss of CHF 8.2 mn. in 2005, the new formula would, ceteris paribus, have caused a total loss of CHF 85.6 mn. $(= 8.2 + 101.6 - 24.2)$.

- For the years 2006 and 2007, payments according to the new RA formula are estimated to be CHF 73.5 and CHF 82.3 mn., respectively, compared to the CHF 2.6 mn. and CHF 2.3 mn. under the current RA formula. In terms of premium income, the shares would have been 9 and 13 percent, respectively, resulting in losses of CHF 54.8 and CHF 86.2 mn., ceteris paribus.

- Payments of A into the RA scheme increase in all cantons. In some, A even turns from receiver into payer, such as in the cantons of Vaud (VD) and Geneva (GE). This precludes a regional restructuring of A's business as a possible RM response; for this reason, this alternative will not be discussed in Section 5.3.3 below.

Arguably, these developments would have jeopardized A's economic survival. Starting with the underwriting result, the combined ratio (defined as loss payments plus

[6]For added precision, calculations are based on months of contract life.

administrative expense plus RA values relative to premium income) was very close to 100 percent over the time period considered, viz. 102.3 (2005), 99.8 (2006), and 100.3 percent (2007).[7] This is not fatal as long as the insurer is making enough profits from capital investment (see e.g. Zweifel and Eisen, 2005, Chapter 5), which was indeed the case in 2007. However, the new RA formula would have caused the combined ratio to attain 111.9 (2005), 107.5 (2006), and 110.7 percent (2007) respectively, amounts that could not have easily been compensated by profits from capital investment. According to Browne and Hoyt (1995), who analyze market predictors of insolvencies in US property-liability insurance between 1970 and 1990, a 5 point increase of the combined ratio causes the insolvency rate to increase by roughly 22 percent. Even if this result cannot be directly applied to health insurers operating in a different country, a 10 point hike in the combined ratio must substantially increase the insolvency risk of an insurer who has limited reserves. The ordinance on health insurance (Federal Council of Switzerland, 2003) requires insurers to hold reserves as a function of enrollment. With more than 150,000 insured, A currently must have reserves amounting to 10 percent of annual premiums (santésuisse, 2009). If A would have used its reserves to make up for the predicted loss of 2005 under the new RA formula, this ratio would have fallen to around 5 percent. The predicted loss of 2006 and 2007 would have wiped out its reserves altogether.

The insolvency of an insurer could be the result of lackluster performance and hence of little importance to the economy as a whole. However, this does not seem to be true of insurer A. It did incur a loss in 2005 but was able to turn this into a surplus for the years 2006 and 2007. In addition, its high predicted payments into RA under the new RA formula are due to its low hospitalization rates (see Figure 5.3). For men (gray bars), they are significantly lower than the Swiss average (black bars) across all age groups (women similar but not shown). While successful risk selection cannot be excluded completely as an explanation, the evidence points in a different direction.

[7]The expense ratio was 5.6 (2005), 5.9 (2006) and 5.6 percent (2007), which is average for Swiss statutory health insurers.

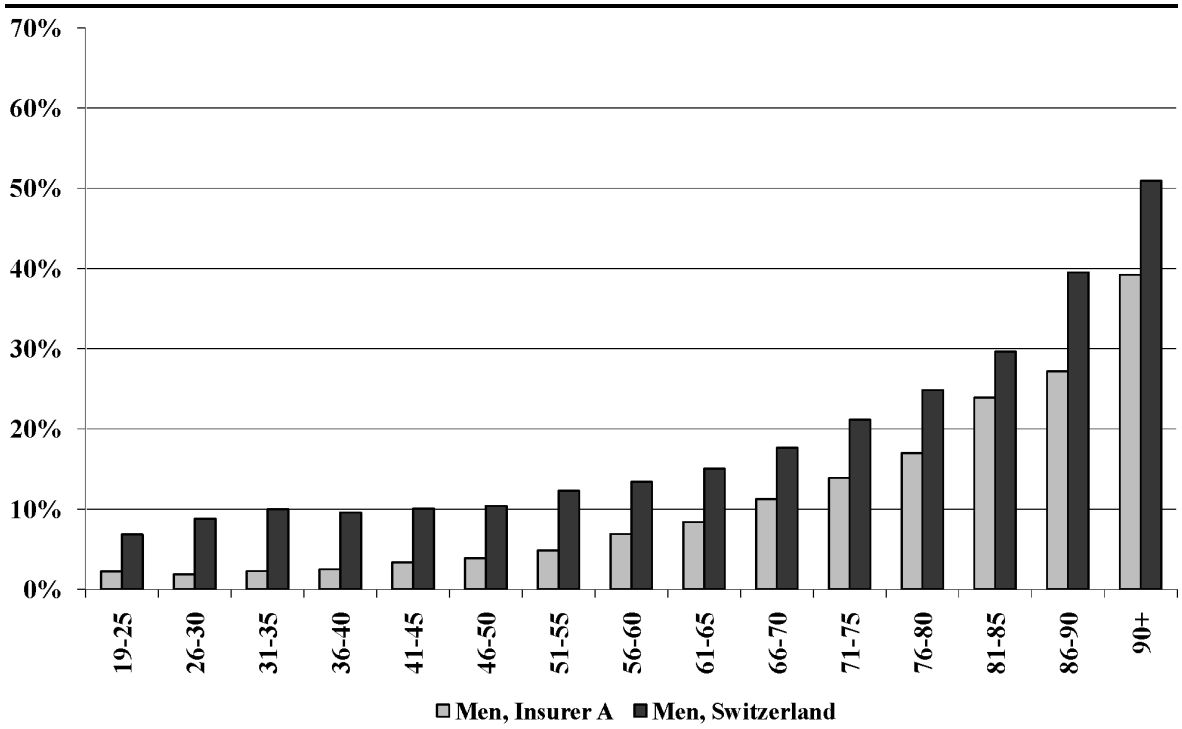


Figure 5.3: Hospitalization rate, insurer A vs. simulated nationwide values, men (2005)

First, as stated in Section 5.2.2, the younger age classes and men are only slightly over-represented. A systematic risk selector would have significantly higher market shares in this age segment. Second, MC contracts (designed to prevent or shorten hospital stays) attain a share of 35 percent in 2007, way above the Swiss average of 16.9 percent. At the same time, insurer A's distribution of MC contracts across age classes does not systematically differ from that of the representative three insurers. Third, total HCE per enrollee and its age profile are quite similar between insurer A and the three others, speaking against across-the-board risk-selection effort on the part of the insurer A. By way of contrast, Figures 5.4 and 5.5 reveal a marked difference with regard to the cost of inpatient and outpatient care. Starting with the age group 51 to 55 but especially beyond age 81, insurer A is markedly below the simulated nationwide benchmark (Figure 5.4). Now this could still be due to risk-selection efforts cleverly targeted at the healthy elderly. In that case, however, one would also expect insurer A's cost for outpatient care to be comparatively low in the higher age groups.

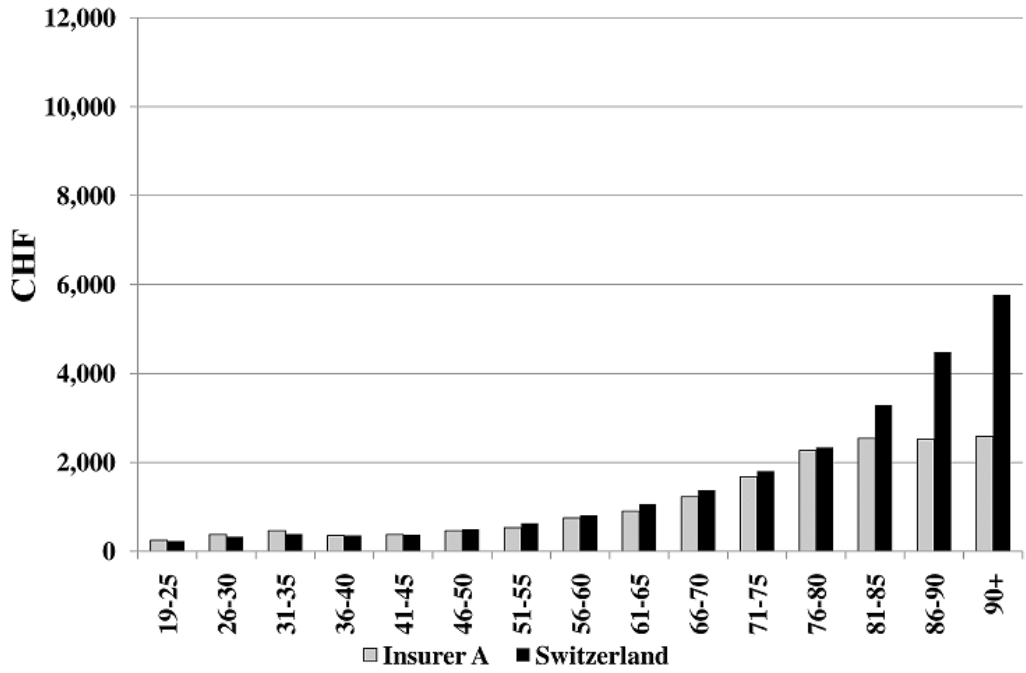


Figure 5.4: Inpatient cost, insurer A vs. simulated nationwide values, CHF (2005)

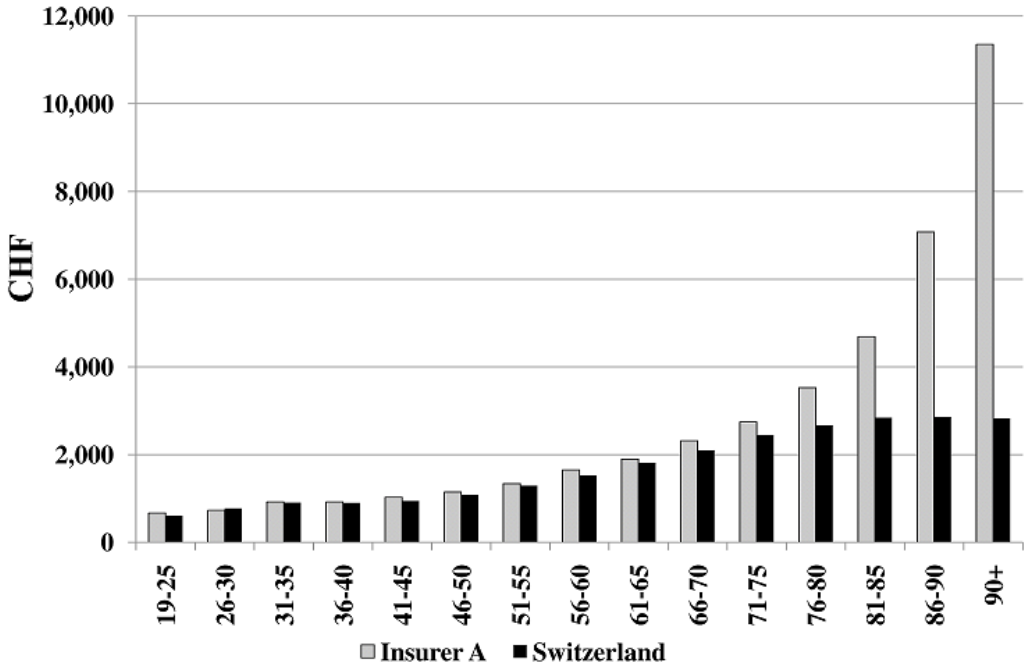


Figure 5.5: Outpatient cost, insurer A vs. simulated nationwide values, CHF (2005)

Yet Figure 5.5 shows that insurer A's cost of outpatient care per enrollee is higher than that of the three representative insurers, and particularly so in the high age groups.

These findings lend credibility to insurer A's claim to have implemented MC in general and home care instead of hospital care specifically for the elderly. This has positive effects not only for the individual patient whose quality of life is higher, but also for the economy as a whole. Indeed, the cost of inpatient care evidenced in Figure 5.4 is only one-half of the true value since the cantons finance roughly 50 percent of hospitals' operating cost. Implementation of MC concepts thus provides relief to taxpayers. Hence, rather than acting as a "cherry-picker", insurer A seems to be among the foremost in conforming with stated objectives of Swiss health policy, i.e. to achieve savings through MC. Insolvency of such an insurer caused by a change in the RA formula can be justifiably qualified as regulatory failure.

5.3.3 Impact on Risk Management

It is unlikely that an insurer confronted with the changes described in the preceding sections can continue with its RM strategy unchanged. The two main alternatives revolve around the two principal activities of an insurer, viz. underwriting and capital investment. Starting with the latter, the insurer could seek offsetting returns on capital investments. However, in the present state of the economy this is very difficult. In addition, capital market theory predicts that higher expected returns can only be achieved in return for more risk once the efficient frontier has been reached, a consequence that is not easily accepted by a regulator of social health insurance. The second possibility is to increase margins from underwriting either by increasing net premiums or reducing claims. Swiss statutory health insurers have to pay by law for all services included in the official list of benefits, with most prices regulated. Therefore, it is not possible to decrease insurance claims significantly. Liabilities arising from underwriting can be reduced by purchasing reinsurance; however, up to present reinsurers have not been providing coverage against RA liabilities. This

leaves an increase of premiums net of RA payments as the likely RM response. Since premiums are fixed by community-rating regulation, lowering payments into the RA scheme becomes the preferred alternative.

One way to achieve this objective is to enroll more unfavorable risks, in particular persons who were hospitalized during the previous year. This is the adjustment the new RA formula was designed to bring about. The challenge to the insurer's RM now becomes to achieve more hospitalizations without incurring much additional cost. Recall that a hospitalization counts as soon as it exceeds three days. When segmenting A's HCE function according to length of stay in the hospital during the previous year, it turns out that patients with four days do not cost significantly more than those with three. Therefore, A has to weigh the once-and-for-all extra cost of a hospital day against the extra contribution from the RA scheme, which may amount to several thousand CHF (see Table 4.2 in Chapter 4, page 78).

The possible reduction of RA payments can be estimated as follows. While it may not be possible to collude with the public hospitals (who obtain a per diem roughly twice the amount paid by the insurer because one-half of their extra operating cost is covered by the canton) to extend all hospital stays from three to four days, this should be possible in 50 percent of all cases. The effect of such a RM response can be estimated with sufficient precision for the three cantons where A has the highest market share (viz. Zurich (ZH), Berne (BE), and Vaud (VD)). There, it would have reduced RA payments by CHF 5 mn. in 2007. Extrapolating to A's entire book of business, one obtains CHF 9 mn., or 11.2 percent of the estimated CHF 82.3 mn. Savings of this magnitude would have been important enough to induce a change in RM.

The cost of this change would fall on taxpayers (who cover one half of the increased operating costs of public hospital through cantonal subsidies), employers (who bear the workdays lost), and patients (who presumably enjoy a higher quality of life outside the hospital). For this reason, reducing the length of hospital stays has been a stated goal

of Swiss health policy, notably justifying the introduction of hospital payment through diagnosis-related groups by 2012 (DRGs, see SwissDRG, 2009). Thus, the fine-tuning of regulation through an improvement of the RA formula risks to burden the economy with sizable inefficiencies.

5.4 Conclusions

Regulation may pose unintended challenges to the risk management (RM) of a company. This chapter analyzes the case of health insurance, where the imposition of community-rating creates an incentive to select favorable risks. Risk adjustment (RA) schemes have been implemented in several countries such as Germany, Israel, the Netherlands, and the United States to counteract this incentive. They make insurers with an above-average share of favorable risks (indicated by age, gender, and other adjusters) to pay into the scheme, which supports insurers with an above-average share of unfavorable risks. Since its current RA formula fails to neutralize the incentive for risk selection, Switzerland will complement it in 2012 with the adjuster, "Hospitalization of more than three days or living in a nursing home during the previous year". This seemingly minor fine-tuning of regulation is shown to have a potentially fatal effect on a particular health insurer A whose payments into the RA scheme would have increased substantially between 2005 and 2007 if the new RA formula had been in effect. The reason is a low rate of hospitalization thanks to a commitment to Managed Care (MC). Therefore, A's most likely RM response would have been to increase recognized hospitalizations by increasing length of stay from three to four days, triggering extra payments from the RA scheme at a limited once-and-for-all cost of an extra hospital day. The cost of this change of RM strategy would have been borne by taxpayers (through increased subsidies of hospitals' operating expense), employers (through lost workdays), and patients (through lower quality of life).

There are lessons to be learned for other countries who impose community-rating on competitive health insurers. First, it is practically impossible to fully neutralize insur-

ers' risk selection incentive through an RA scheme,[8] and be it only due to their different rates of discount in estimating the present value of the benefits and costs associated with risk selection. Second, perfecting the RA formula can have unintended side effects at the level of an individual insurer that go as far as jeopardizing its economic survival in spite of innovative effort. In the case studied here, the insurer is even punished for its innovative commitment to MC. Finally, the threat of survival may well trigger adjustments in RM strategy that cause an efficiency loss to the economy as a whole.

Acknowledgments

The authors gratefully acknowledge helpful suggestions by Maria Trottmann (University of Zurich, Switzerland), Frank Lichtenberg (Columbia University, New York, USA), and an anonymous referee. Furthermore, special thanks go to the three health insurers and health insurer A that provided data for this study.

[8]There is the perception that for all its refinement, the all-encounter RA CMS-HCC model overpays Medicare Advantage Programs (representing MCOs). We owe this interesting point to the anonymous referee.

Chapter VI
Conclusions

6 Conclusions

This chapter discusses policy implications and possible extensions of each essay.

If premiums are community-rated in social statutory health insurance, preferences of the insured have to be the crucial factor for determining the uniform benefit package. From a cost-benefit perspective and neglecting distributional concerns, inclusion of a treatment or pharmaceutical is justified if the insured exhibit a willingness-to-pay that exceeds the cost of treatment, i.e. not only patients' (ex-post) willingness-to-pay, but also potential patients' (ex-ante) willingness-to-pay has to be sufficiently high. Chapter 2 demonstrates that this indeed seems to be the case for the long-acting insulin analogue "Insulin Detemir" in Germany. Respondents not only value attributes such as avoidance of hypoglycemia and weight gain, but also attach importance to attributes that typically are judged medically irrelevant, such as lack of preparation (swinging) the insulin before use and flexibility of timing of the injection. For these reasons, the new insulin analogue should be included in the German statutory health insurance list of benefits.

The implications of Chapter 3 appear to be very specific, but address an issue of general importance. In particular, the results show that the assumption of a linearly approximated utility function can result in biased estimates of willingness-to-pay. In general, however, Chapter 3 illustrates that model assumptions can have significant impact on produced results. Hence, when performing discrete-choice experiments, the underlying assumptions need to be tested. When making decisions based on discrete-choice experiments, one has to be aware that the results rely on assumptions. This is not only an issue solely in such experiments or in health economics. For

example, this problem surfaced during the financial credit risk management crisis of the last three years, when the Gaussian-copula assumption in the pricing of credit derivatives led to underestimation of extreme events (see Donnelly and Embrechts, 2010).

The findings of Chapter 4 outline that policy makers need to distinguish between the insurers, who pay contributions to the risk-adjustment scheme, and the favorable risks, who ultimately bear these cross-subsidies by paying more than their actuarially fair premium. With the criteria of age and gender, the elderly mainly receive payments from risk adjustment. Including the additional criterion "Hospitalization or living in a nursing home during the previous year" ("hospitalization" henceforth) changes the redistribution profile significantly. Hospitalization becomes the crucial factor. When adding additional risk adjusters, such as diagnostic categories or pharmaceutical cost groups, the redistribution effects have to be reexamined. However, there is a trade-off when adding hospitalization to the risk-adjustment formula. This refinement reduces insurers' incentives for risk selection, but it will also increasingly shelter insurers from financial risk, undermining efficiency. Consequently, a cap on the volume of risk adjustment may well improve insurers' efficiency, but it also increases incentives for risk selection. To minimize this cost, Chapter 4 proposes a uniform reduction of positive and negative risk-adjustment values by a certain amount, depending on existing differences between groups in terms of risk-adjustment values and their population shares.

The inclusion of the hospitalization criterion in the risk-adjustment formula was expected to punish only the "cherry-pickers" among the health insurers. However, Chapter 5 illustrates the unintended side effects for an innovative insurer who is among the foremost in conforming to the stated objectives of Swiss health policy, i.e. to achieve savings through Managed Care. This can go as far as to jeopardize the firm's economic survival, punishing it for its innovative effort. This threat to survival may trigger adjustments in risk-management strategy that cause efficiency losses to the economy as a whole. For this reason, policy makers should consider the direct and

partial effects of regulations alongside their side effects and general implications.

Each chapter could form the basis for the extension of future research.

The survey conducted in Chapter 2 includes much information that has not yet been evaluated. So far, only heterogeneity with respect to diabetic status has been analyzed. Further criteria for heterogeneity could be discovered. Conversely, it would also be interesting to analyze homogeneity observed in certain cases. For example, non-diabetics were asked if one of their closest family members was diabetic. One would expect these respondents to have a higher willingness-to-pay for modern insulin therapy; presumably they would be altruistic toward their family members, but their own risk to become insulin-dependent might be above average, too (Hauner, 2010). Another extension of Chapter 2 could be to apply the findings of Chapter 3. The discrete-choice experiment of Chapter 2 could be re-estimated using a non-linear utility function.

Chapter 3 addresses just one specific issue of the discrete-choice approach. However, around 40 percent of respondents in this experiment never deviated from their status quo. No information from these respondents could be gained because they never "jump" over the indifference curve. Future research could model this decision process by a two-stage estimation procedure, similar to the two-part model used to estimate health care expenditure. In a first step, one would estimate the probability that the respondent chooses the alternative at least once. The second step would model the probability of choosing the alternative, given that the respondent deviated from the status quo at least once.

The most likely future development in Swiss risk adjustment comes from the formula being extended by further criteria. A likely candidate is pharmaceutical cost groups, as already in use in the Netherlands (Lamers and Van Vliet, 2003b). Beck et al. (2010) found that these groups would decrease incentives for risk selection significantly

in Switzerland. An enhancement of Chapter 4 would simulate the effects of this additional criterion on the volume of cross-subsidization. However, the required individual-level data is not yet accessible. Another future research question could utilize more sophisticated risk measures in Chapter 4. Instead of the variance, either value-at-risk or expected shortfall would disclose the optimization of the derived uniform reductions per risk class.

While the hospitalization criterion successfully weakens incentives for risk selection, insurers are only compensated for inpatient treatments. Adding diagnostic or pharmaceutical cost groups into the risk-adjustment formula could mitigate the resulting incentives. If such individual-level data were available, the consequences for insurer A could be simulated. Payments into the risk-adjustment scheme would probably be much lower, accounting for the insurer's higher-than-average outpatient costs.

However, there is an alternative that avoids the issues induced by community-rating. According to Rothschild and Stiglitz (1976) the first-best Pareto-optimum equilibrium for (health) insurance, given public information about risk, is full coverage for both high and low risks, associated with high and low premiums. That is, health insurers would be permitted to charge premiums reflecting the difference in risk. With sufficient pressure of competition, this would boil down to "price equal to expected marginal cost", since expected future health care expenditure importantly reflect the insurer's cost of enrolling an additional customer. However, when introducing risk-based premiums, equity issues would likely be raised. Wealthy individuals could pay a high risk-based premium out of their own pockets. Low-income individuals who are favorable risks could also pay for their own premiums. Low-income individuals who are unfavorable risks, however, are a problematic group. For this reason, instead of community-rating, an additional solution requiring less market intervention has been proposed in the literature. Low-income unfavorable risks would be entitled to an earmarked subsidy that kicks in as soon as their premium exceeds a certain percentage of their income (see Zweifel and Breuer, 2006). In fact, the new health insurance law

of 2004 introduced such a targeted subsidy in Switzerland without eliminating the premium regulation introduced in 1911. It is thus unlikely that risk-based premiums will be introduced in Switzerland soon. For this reason, the topic of this dissertation will remain important in the future.

Bibliography

Amaya-Amaya, M., Gerard, K., Ryan, M., 2008. Discrete choice experiments in a nutshell. In: Ryan, M., Gerard, K., Amaya-Amaya, M. (Eds.), Using Discrete Choice Experiments to Value Health and Health Care. Springer, Dordrecht, Ch. 1, pp. 13–46.

American Diabetes Association, 2010. Diabetes basics. Website.
URL `www.diabetes.org/diabetes-basics/`

Aristides, M., Weston, A. R., FitzGerald, P., Le Reun, C., Maniadakis, N., 2004. Patient preference and willingness-to-pay for humalog mix25 relative to humulin 30/70: a multicountry application of a discrete choice experiment. Value In Health: The Journal of The International Society for Pharmacoeconomics and Outcomes Research 7 (4), 442 – 454.

Atkinson, A., Donev, A., 1992. Optimum Experimental Designs. Claredon Press, Oxford.

Basu, A., Bhakti, V. A., Rathour, P. J., 2006. Scale of interest versus scale of estimation: Comparing alternative estimators for the incremental costs of a comorbidity. Health Economics 15, 1091–1107.

Basu, A., Manning, W. G., Mullahy, J., 2004. Comparing alternative models: log vs cox proportional hazard. Health Economics 13, 749–765.

Beck, K., 2004. Risiko Krankenversicherung. Risikomanagement in einem regulierten Krankenversicherungsmarkt - (Risk of Health Insurance. Risk Management in Regulated Health Insurance). Haupt Verlag, Bern.

Beck, K., Trottmann, M., Käser, U., Keller, B., von Rotz, S., Zweifel, P., 2006. Nachhaltige Gestaltung des Risikoausgleichs in der Schweizer Krankenversicherung - (Sustainable Design of Swiss Risk Adjustment). Ott Verlag, Basel.

Beck, K., Trottmann, M., Zweifel, P., 2010. Risk adjustment in health insurance and its long-term effectiveness. Journal of Health Economics 29 (4), 489–498.

Becker, K., 2006. Flexibilisierungsmöglichkeiten in der Krankenversicherung - (Possibilities of more Flexibility in Health Insurance). Dr. Kovac Verlag, Hamburg.

Becker, K., Brändle, A., Zweifel, P., 2007. Was wollen die Versicherten? Evidenz aus Deutschland und den Niederlanden - (What do the Insured Want? Evidence from Germany and the Netherlands). Bertelsmann Stiftung, Gütersloh.

Becker, K., Zweifel, P., 2008. Age and choice in health insurance: Evidence from a discrete-choice experiment. The Patient: Patient-Centered Outcome Research 1, 27–40.

Bindman, A. B., Chattopadhyay, A., Osmond, D. H., Huen, W., Bacchetti, P., 2005. The impact of medicaid managed care on hospitalizations for ambulatory care sensitive conditions. Health Services Research 40 (1), 19 – 37.

Browne, M. J., Hoyt, R. E., 1995. Economic and market predictors of insolvencies in the property-liability industry. The Journal of Risk and Insurance 62 (2), 309–327.

Burgess, L., Street, D., 2003. Optimal designs for 2^k choice experiments. Communications in Statistics: Theory and Methods 32, 2185–2206.

Bush, M. A., 2007. Intensive diabetes therapy and body weight: focus on insulin detemir. Endocrinology and Metabolism Clinics of North America 36 Suppl 1, 33 – 44.

Calfo, S., 2009. Medicare risk adjustment. Website, center for Medicare and Medicaid Services.
URL `www.maac-actuary.org/Past_Meetings/2008_Annual_Meeting/Session_5B_Medicare_Risk_Adjustment.ppt`

Cameron, C., Bennett, H., 2009. Cost-effectiveness of insulin analogues for diabetes mellitus. Canadian Medical Association Journal 180 (4), 400–407.

Carlsson, F., Martinsson, P., 2003. Design techniques for stated preference methods in health economics. Health Economics 12, 281–294.

Cleveland, W., 1979. Robust locally weighted regression and smoothing scatterplots. Journal of the American Statistical Association 47, 829–836.

Cleveland, W., Devlin, S., 1988. Locally weighted regression: An approach to regression analysis by local fitting. Journal of the American Statistical Association 83, 596–610.

Copas, J., 1983. Regression, prediction and shrinkage (with discussion). Journal of the Royal Statistical Society B 45, 311–354.

Culyer, A., 1990. Commodities, characteristics of commodities, characteristics of people, utilities, and quality of life. In: Baldwin, S., Godfrey, C., Propper, C. (Eds.), Quality of Life: Perspectives and Policies. Routledge, London, pp. 9–27.

Davey, P., Grainger, D., MacMillan, J., Rajan, N., Aristides, M., Dobson, M., 1998. Economic evaluation of insulin lispro versus neutral (regular) insulin therapy using a willingness-to-pay approach. Pharmacoeconomics 13 (3), 347–358.

Demssie, Y. N., Younis, N., Soran, H., 2009. The role of insulin detemir in overweight type 2 diabetes management. Vascular Health and Risk Management 5, 553–560.

Donnelly, C., Embrechts, P., 2010. The devil is in the tails: actuarial mathematics and the subprime mortgage crisis. ASTIN Bulletin 40 (1), 1–33.

Dornhorst, A., Lüddeke, H.-J., Honka, M., Ackermann, R. W., Meriläinen, M., Gallwitz, B., Sreenan, S., 2008. Safety and efficacy of insulin detemir basal-bolus therapy in type 1 diabetes patients: 14-week data from the European cohort of the PREDICTIVE study. Current Medical Research And Opinion 24 (2), 369 – 376.

Douven, R., 2007. Morbidity-based risk adjustment in the Netherlands. In: Wille, V., Schneider, U. (Eds.), Wettbewerb und Risikostrukturausgleich im internationalen Vergleich; Erfahrungen aus den USA, der Schweiz, den Niederlanden und Deutschland - (Competition and risk adjustment; experiences from the US, Switzerland, the Netherlands, and Germany). Nomos, pp. 161–202.

Dranitsaris, G., Longo, C. J., Grossman, L. D., 2000. The economic value of a new insulin preparation, humalog mix 25: Measured by a willingness-to-pay approach. Pharmacoeconomics 18 (3), 275–287.

Drummond, M., Sculpher, M., Torrance, G., O'Brien, B., Stoddart, G., 2005. Methods for the Economic Evaluation of Health Care Programmes, 3rd Edition. Oxford University Press, Oxford.

Efron, B., Tibshirani, R., 1993. An introduction to the bootstrap. Chapman & Hall, New York.

Ellis, R. P., Van de Ven, W. P., 2000. Risk adjustment in competitive health plan markets. In: Culyer, A., Newhouse, J. (Eds.), Handbook of Health Economics. North-Holland, Amsterdam, pp. 755–845.

Enthoven, A., 1986. Managed competition in health care and the unfinished agenda. Health Care Financing Review Annual Supplement (7), 105–120.

Federal Council of Switzerland, 2003. Verordnung über die Krankenversicherung - (Regulation on health insurance). Website.
URL `http://www.admin.ch/ch/d/sr/832_102/`

Federal Ministry of Health, 2008. Bundesministerium für Gesundheit Deutschland - (The Federal Ministry of Health Germany). Website.
URL `www.die-gesundheitsreform.de`

Federal Office of Public Health, 2005. Statistik der obligatorischen Krankenpflegeversicherung - (Swiss Statutory Health Insurance Statistics). Bern.

Federal Office of Public Health, 2007. Statistik der obligatorischen Krankenpflegeversicherung - (Swiss Statutory Health Insurance Statistics). Bern.

Fieller, E., 1954. Some problems in interval estimation with discussion. Journal of the Royal Statistical Society 16, 175–185.

Freeman, J., 2009. Insulin analog therapy: Improving the match with physiologic insulin secretion. The Journal of the American Osteopathic Association 109 (1), 26–36.

Gerard, K., Shanahan, M., Louviere, J., 2008. Using discrete choice modeling to investigate breast screening participation. In: Ryan, M., Gerard, K., Amaya-Amaya, M. (Eds.), Using Discrete Choice Experiments to Value Health and Health Care. Springer, Dordrecht, Ch. 5, pp. 117–137.

GFS, 2001. Qualitäts- und Kostenorientierung, Trendstudie zum Gesundheitsmonitor - (Quality- and Costorientation, Trendstudy for Monitoring Health). Interpharma, Basel.

Giani, G., Janka, H., Hauner, H., Standl, E., Schiel, R., Neu, A., Rathmann, W., Rosenbauer, J., 2004. Epidemiologie und Verlauf des Diabetes mellitus in Deutschland - (Epidemiology and development of diabetes mellitus in Germany). Website.
URL `www.deutsche-diabetes-gesellschaft.de`

Glazer, J., McGuire, T. G., 2002. Setting health plan premiums to ensure efficient quality in health care: Minimum variance optimal risk adjustment. Journal of Public Economics 84, 153–173.

Gleser, L. J., Hwang, J. T., 1987. The nonexistence of 100(1 - α)% confidence sets of finite expected diameter in errors-in-variables and related models. The Annals of Statistics 15 (3), 1351–1362.

Green, C., Gerard, K., 2009. Exploring the social value of health-care interventions: A stated preference discrete choice experiment. Health Economics 18, 951–973.

Gschwend, M. H., Aagren, M., Valentine, W. J., 2009. Cost-effectiveness of insulin detemir compared with neutral protamine Hagedorn insulin in patients with type 1 diabetes using a basal-bolus regimen in five European countries. Journal of Medical Economics 12 (2), 114–123.

Guimarães, C., Marra, C. A., Colley, L., Gill, S., Simpson, S., Meneilly, G., Queiroz, R. H. C., Lynd, L. D., 2009a. Socioeconomic differences in preferences and willingness-

to-pay for insulin delivery systems in type 1 and type 2 diabetes. Diabetes Technology and Therapeutics 11 (9), 567 – 573.

Guimarães, C., Marra, C. A., Colley, L., Gill, S., Simpson, S. H., Meneilly, G. S., Queiroz, R. H. C., Lynd, L. D., 2009b. A valuation of patients' willingness-to-pay for insulin delivery in diabetes. International Journal of Technology Assessment in Health Care 25 (3), 359 – 366.

Haak, T., Tiengo, A., Draeger, E., Suntum, M., Waldhäusl, W., 2005. Lower within-subject variability of fasting blood glucose and reduced weight gain with insulin detemir compared to NPH insulin in patients with type 2 diabetes. Diabetes, Obesity and Metabolism 7, 56–64.

Haak, T., Tiengo, A., Waldhäusl, W., Draeger, E., 2003. Treatment with insulin detemir is associated with predictable fasting blood glucose levels and favourable weight development in subjects with type 2 diabetes. Diabetes 52 (Suppl.1), A120.

Hanemann, M. W., 1983. Marginal welfare measures for discrete choice models. Economics Letters 13, 129–136.

Harrell, F., Lee, K., Mark, D., 1996. Tutorial in biostatistics: Multivariable prognostic models: Issues in developing models, evaluating assumptions and measuring and reducing errors. Statistics in Medicine 15, 361–387.

Hauber, A. B., Mohamed, A. F., Johnson, F. R., Falvey, H., 2009. Treatment preferences and medication adherence of people with type 2 diabetes using oral glucose-lowering agents. Diabetic Medicine: A Journal of The British Diabetic Association 26 (4), 416 – 424.

Hauner, H., 2008. Diabetesepidemiologie und Dunkelziffer - (Epidemiology and iceberg phenomenon in diabetes). In: Nuder, G. (Ed.), Deutscher Gesundheitsbericht Diabetes 2008. Deutsche Diabetes Union DDU, pp. 7–11.

Hauner, H., 2010. Diabetesepidemiologie und Dunkelziffer - (Epidemiology and iceberg phenomenon in diabetes). In: Nuber, G. (Ed.), Deutscher Gesundheitsbericht Diabetes 2011. DiabetesDE, Berlin, pp. 8–13.

Heitjan, D. F., 2000. Fieller's method and net health benefits. Health Economics 9, 327–335.

Hermansen, K., Davies, M., 2007. Does insulin detemir have a role in reducing risk of insulin-associated weight gain? Diabetes, Obesity and Metabolism 9, 209–217.

Hermansen, K., Dornhorst, A., Sreenan, S., 2009. Observational, open-label study of type 1 and type 2 diabetes patients switching from human insulin to insulin analogue basal-bolus regimens: insights from the PREDICTIVE study. Current Medical Research and Opinion 25 (11), 2601–2608.

Hermansen, K., Fontaine, P., Kukolja, K., 2004. Insulin analogues (insulin detemir and insulin aspart) versus traditional human insulins (NPH insulin and regular human insulin) in basal-bolus therapy for patients with type 1 diabetes. Diabetologia 47, 622–629.

Holly, A., Gardiol, L., Eggli, Y., Yalcin, T., 2003. Health-based risk adjustment in Switzerland: An exploration using medical information from prior hospitalization. Mimeo; Institut d'Economie et Management de la Santé; Lausanne.

Home, P., Bartley, P., Russell-Jones, D., 2004. Insulin detemir offers improved glycemic control compared to NPH insulin in people with type 1 diabetes: A randomized clinical trial. Diabetes Care 27, 1081–1087.

Horvath, K., Jeitler, K., Berghold, A., Ebrahim, S., Gratzer, T., Plank, J., Kaiser, T., Pieber, T., Siebenhofer, A., 2007. Long-acting insulin analogues versus NPH insulin (human isophane insulin) for type 2 diabetes mellitus. Cochrane Database of Systematic Reviews 2 (CD005613).

Hosmer, D., Lemeshow, S., 1980. A goodness-of-fit test for the multiple logistic regression model. Communications in Statistics A10, 1043–1069.

Hosmer, D., Lemeshow, S., 2000. Applied Logistic Regression, 2nd Edition. Wiley, New York.

Hoyos, D., 2010. The state of the art of environmental valuation with discrete choice experiments. Ecological Economics, The Transdisciplinary Journal of the International Society for Ecological Economics 69 (8), 1595–1603.

Huang, E. S., Basu, A., O'Grady, M., Capretta, J. C., 2009. Projecting the future diabetes population size and related costs for the U.S. Diabetes Care 32 (12), 2225 – 2229.

Hull, J. C., 2006. Options, Futures and Other Derivatives, 6th Edition. Pearson Prentice Hall, New Jersey.

Häussler, B., Berger, U., Mast, O., Thefeld, W., 2005. Risk and potential risk reduction in diabetes type 2 patients in Germany. European Journal of Health Economics 6 (2), 152–158.

IQWiG, 2009. Langwirksame Insulinanaloga zur Behandlung des Diabetes mellitus Typ 2 – Abschlussbericht - (Long-acting insulin analogues for treatment of diabetes mellitus type 2 - Final report). Version 1.1 A05-03, Institute for Quality and Efficiency in Health Care.

IQWiG, 2010. Langwirksame Insulinanaloga zur Behandlung des Diabetes mellitus Typ 1 – Abschlussbericht - (Long-acting insulin analogues for treatment of diabetes mellitus type 1 - Final report). Version 1.0 A05-01, Institute for Quality and Efficiency in Health Care.

Jack, W., 2006. Optimal risk adjustment with adverse selection and spatial competition. Journal of Health Economics 25, 908–926.

Johnson, F., Desvousges, W., 1997. Estimating stated preferences with rated-pair data: Environmental, health, and employment effects of energy programs. Journal of Environmental Economics and Management 34, 79–99.

Joint Organization KVG, 2005. Geschäftsbericht - (Business report).

Joint Organization KVG, 2008. Geschäftsbericht - (Business report).

Kennedy, P. E., 2003. A Guide to Econometrics, 5th Edition. MIT Press, Cambridge.

Kjaer, T., Bech, M., Gyrd-Hansen, D., Hart-Hansen, K., 2006. Ordering effect and price sensitivity in discrete choice experiments: Need we worry? Health Economics 15 (11), 1217 – 1228.

Kolendorf, K., Pavlic-Renar, I., Santeusanio, F., A., P., Gall, M., Heller, S., 2004. Insulin detemir is associated with lower risk of hypoglycemia compared to NPH insulin in people with type 1 diabetes. Program of American Diabetes Association's 64th annual scientific sessions, A551–P.

Kominski, G. F., 2007. Medicare's use of risk adjustment. National Health Policy Forum.

Kuhfeld, W., Tobias, R., Garratt, M., 1994. Efficient experimental design with marketing research applications. Journal of Marketing Research XXXI, 545–557.

Kurtzhals, P., 2007. Pharmacology of insulin detemir. Endocrinology and Metabolism Clinics of North America 36 (Suppl. S1), 6–52.

Lamers, L. M., 1999. Risk adjusted capitation based on the diagnostic cost group model: An empirical evaluation with health survey information. Health Services Research 33 (6), 1727–44.

Lamers, L. M., Van Vliet, R. C. J. A., 2003a. Health-based risk adjustment: Improving the pharmacy-based cost group model to reduce gaming possibilities. European Journal of Health Economics 4, 107–114.

Lamers, L. M., Van Vliet, R. C. J. A., 2003b. The Pharmacy-based Cost Group Model: Validating and adjusting the classification of medications for chronic conditions to the Dutch situation. Health Policy 68, 113–128.

Lancsar, E., Louviere, J., 2006. Deleting "irrational" respondents from discrete choice experiments: A case of investigating or imposing preferences? Health Economics 15, 797–811.

Lehmann, H., Zweifel, P., 2004. Innovation and risk selection in deregulated social health insurance. Journal of Health Economics 23, 997–1012.

Leichter, S., 2008. Is the use of insulin analogue cost-effective? Advances in Therapy 25 (4), 285–299.

Louviere, J. J., Hensher, D. A., Swait, J. D., 2000. Stated Choice Methods - Analysis and Application. Cambridge University Press, Cambridge.

Louviere, J. J., Lancsar, E., 2009. Choice experiments in health: The good, the bad, the ugly and toward a brighter future. Health Economics, Policy and Law 4 (4), 527–546.

Luce, D., 1959. Individual Choice Behavior. Wiley and Sons, New York.

Mandosi, E., Fallarino, M., Rossetti, M., Gatti, A., Morano, S., 2009. Waist circumference reduction after insulin detemir therapy in type 2 diabetes patients previously treated with NPH. Diabetes Research And Clinical Practice 84 (2), e18 – e20.

Manski, C., Lerman, S. R., 1977. The estimation of choice probabilities from choice based samples. Econometrica 45 (8), 1977–1988.

Marre, M., Pinget, M., Gin, H., Thivolet, C., Hanaire, H., Robert, J., Fontaine, P., 2009. Insulin detemir improves glycaemic control with less hypoglycaemia and no weight gain: 52-week data from the PREDICTIVE study in a cohort of French patients with type 1 or type 2 diabetes. Diabetes & Metabolism 35 (6), 469–475.

McFadden, D., 1974. Conditional logit analysis of qualitative choice behavior. In: Zarembka, P. (Ed.), Frontiers of Econometrics. Academic Press, New York, pp. 105–142.

McFadden, D., 1981. Econometric models of probabilistic choice. In: Manski, C., McFadden, D. (Eds.), Structural Analysis of Discrete Data with Econometric Applications. The MIT Press, Cambridge, pp. 198–272.

McFadden, D., 2001. Economic choices. The American Economic Review 91 (3), 351–378.

McNiel, A. J., Frey, R., Embrechts, P., 2005. Quantitative Risk Management: Concepts, Techniques, and Tools. Princeton University Press, Princeton.

Mickey, J., Greenland, S., 1989. A study of the impact of confounder-selection criteria on effect estimation. American Journal of Epidemiology 129, 125–137.

Miller, R., Luft, H., 1997. Does managed care lead to better or worse quality of care? Health Affairs 16 (5), 7–25.

Monami, M., Marchionni, N., Mannucci, E., 2009. Long-acting insulin analogues vs. NPH human insulin in type 1 diabetics. A meta-analysis. Diabetes, Obesity & Metabolism 11 (4), 372–378.

Mullahy, J., Manning, W. G., 1996. Statistical issues in cost-effectiveness analysis. In: Sloan, F. A. (Ed.), Valuing Health Care - Costs, Benefits, and Effectiveness of Pharmaceuticals and Other Medical Technologies. Cambridge University Press, New York, pp. 149–184.

OECD, 2004. The OECD Health Project: Toward high-performing health systems. OECD, Paris.

OECD, 2010. Oecd.stat extracts. Website.
URL `stats.oecd.org`

Palmer, A., Lammert, M., Hermansen, K., 2008. Health economic consequences of insulin analogues in the treatment of type 1 diabetes in Denmark. Ugeskrift For Laeger 170 (15), 1250–1254.

Palmer, A., Valentine, W. J., Ray, J. A., Foos, V., Lurati, F., Smith, I., Lammert, M., Roze, S., 2007. An economic assessment of analogue basal-bolus insulin versus human basal-bolus insulin in subjects with type 1 diabetes in the UK. Current Medical Research and Opinion 23 (4), 895–901(7).

Palmer, A. J., Roze, S., Valentine, W. J., Smith, I., Wittrup-Jensen, K. U., 2004. Cost-effectiveness of detemir-based basal/bonus therapy versus NPH-based basal/bolus therapy for type 1 diabetes in a UK setting: An economic analysis based on meta-analysis results of four clinical trials. Current Medical Research and Opinion 20 (11), 1729–1746.

Pauly, M., 1984. Is cream skimming a problem for the competitive medical market? Journal of Health Economics 3, 87–95.

Pope, G. C., Ellis, R. P., Ash, A. S., Liu, C.-F., Ayanian, J. Z., Bates, D. W., Burstin, H., Iezzoni, L. I., Ingber, M., 2000. Principal inpatient diagnostic cost group model for Medicare risk adjustment. Health Care Financing Review 21 (3), 93–118.

Pope, G. C., Kautter, J., Ellis, R. P., Ash, A. S., Ayanian, J. Z., Iezzoni, L. I., Ingber, M., Levy, J., Robst, J., 2004. Risk adjustment of Medicare capitation payments using the CMS-HCC model. Health Care Financing Review 25 (4), 119–141.

Pregibon, D., 1980. Goodness of link tests for generalized linear models. Applied Statistics 29, 15–24.

Pregibon, D., 1981. Logistic regression diagnostics. The Annuals of Statistics 9, 705–724.

Ramsey, J., 1969. Tests for specification errors in classical linear least squares regression analysis. Journal of the Royal Statistical Society (Series B) 31, 350–371.

Raskin, P., 2007. Efficacy and safety of insulin detemir. Endocrinology and Metabolism Clinics of North America 36 Suppl 1, 21 – 32.

Raslová, K., Tamer, S. C., Clauson, P., Karl, D., 2007. Insulin detemir results in less weight gain than NPH insulin when used in basal-bolus therapy for type 2 diabetes mellitus, and this advantage increases with baseline body mass index. Clinical Drug Investigation 27 (4), 279 – 285.

Regier, D. A., Ryan, M., Phimister, E., Marra, C. A., 2009. Bayesian and classical estimation of mixed logit: An application to genetic testing. Journal of Health Economics 28 (3), 598 – 610.

Robertson, K., Schonle, E., Gucev, Z., 2004. Benefits of insulin detemir over NPH insulin in children and adolescents with type 1 diabetes: Lower and more predictable fasting plasma glucose and lower risk of nocturnal hypoglycemia. Program of American Diabetes Association's 64th annual scientific sessions, A606–P.

Rosenbauer, J., Stahl, A., 2010. Häufigkeit des Diabetes mellitus im Kindes- und Jugendalter in Deutschland - (Prevalence of diabetes mellitus in children and young adults in Germany). Diabetologe 6, 177–189.

Rothschild, M., Stiglitz, J., 1976. Equilibrium in competitive insurance markets: An essay on the economics of imperfect information. The Quarterly Journal of Economics 90 (4), 629–649.

Russell-Jones, D., 2007. Insulin detemir and basal insulin therapy. Endocrinology and Metabolism Clinics of North America 36 (Suppl. S1), 6–52.

Russell-Jones, D., Boliner, J., Simpson, R., 2004. Lower and more predictable fasting glucose and reduced risk of nocturnal hypoglycaemia with once daily insulin detemir versus NPH in subjects with type 1 diabetes. Diabetologia 45 (Suppl. 2), A51.

Russell-Jones, D., Khan, R., 2007. Insulin-associated weight gain in diabetes - causes effects and coping strategies. Diabetes, Obesity and Metabolism 9, 799–812.

Ryan, M., Bate, A., 2001. Testing the assumptions of rationality, continuity and symmetry when applying discrete choice experiments in health care. Applied Economics Letters 8, 59–63.

Ryan, M., Watson, V., 2008. Practical issues in conducting a discrete choice experiment. In: Ryan, M., Gerard, K., Amaya-Amaya, M. (Eds.), Using Discrete Choice Experiments to Value Health and Health Care. Springer, Dordrecht, Ch. 3, pp. 73–97.

Sadri, H., MacKeigan, L. D., Leiter, L. A., Einarson, T. R., 2005. Willingness to pay for inhaled insulin: A contingent valuation approach. Pharmacoeconomics 23 (12), 1215–1227.

Salkeld, G., Ryan, M., Short, L., 2000. The veil of experience: Do consumers prefer what they know best? Health Economics 9 (3), 267 – 270.

santésuisse, 2009. Die Reserven in der obligatorischen Krankenpflegeversicherung - (Reserves in statutory health insurance). Website, Swiss Association of Health Insurers. URL http://www.santesuisse.ch/de/

santésuisse, 2010a. Von den Erwachsenen gewählte Franchisestufen 1999-2008 - (Deductibles of the adult insured 1999-2008). Website, santésuisse Datapool. URL http://www.santesuisse.ch/de/dyn_output_graphic_detail.html?content.cdid=28259

santésuisse, 2010b. Von den Versicherten gewählte Modelle mit eingeschränkter Arztwahl 1999-2008 - (Managed care contracts 1999-2008). Website, santésuisse Datapool. URL http://www.santesuisse.ch/de/dyn_output_graphic_detail.html?content.cdid=28268&navid=416

SAS Institute Corp., 1999. Insight user's guide, version 8.

Satish, K. G., Ramachandra, G. N., 2008. Long-acting insulin analogs versus human insulins. Diabetes Technology & Therapeutics 10 (5), 331–332.

Schleser-Mohr, S., 2007. Einfach gut leben - mit Insulin! - (Simply have a good life - using insulin!). Website. URL http://www.humanmedizin-goettingen.de/media/global/tag_der_medizin/tdm_2004_leben_m_diabetes.pdf

Schmeisl, G.-W., 2009. Schulungsbuch für Diabetiker - (Book of Instructions for Diabetics), 6th Edition. Urban & Fischer, Munich.

Shen, J., Ellis, R. P., 2002a. Cost-minimizing risk adjustment. Journal of Health Economics 21 (3), 515–530.

Shen, J., Ellis, R. P., 2002b. How profitable is risk-selection? A comparison of four risk adjustment models. Health Economics 11, 165–174.

Singh, S. R., Ahmad, F., Lal, A., Yu, C., Bai, Z., Bennett, H., 2009. Efficacy and safety of insulin analogues for the management of diabetes mellitus: A meta-analysis. Canadian Medical Association Journal 180 (4), 385–397.

Skjoldborg, U. S., Gyrd-Hansen, D., 2003. Conjoint analysis. The cost variable: An Achilles' heel? Health Economics 12, 479–491.

Sloane, N., Hardin, R., 2007. Gosset: A general-purpose program for designing experiments. Website.
URL `www.research.att.com`

Soran, H., Younis, N., 2006. Insulin detemir: A new insulin analogue. Diabetes, Obesity and Metabolism 8, 26–30.

Spycher, S., 2000. Reform des Risikoausgleichs in der Krankenversicherung - (Risk Adjustment Reforms in Health Insurance). In: Bundesamt für Sozialversicherung (Ed.), Beiträge zur sozialen Sicherheit. Berne.

Sreenan, S., Virkamäki, A., Zhang, K., Hansen, J., 2008. Switching from NPH insulin to once-daily insulin detemir in basal-bolus-treated patients with diabetes mellitus: Data from the European cohort of the PREDICTIVE study. International Journal of Clinical Practice 62 (12), 1971–1980.

Statistical Offices of the Länder, 2009. Volkswirtschaftliche Gesamtrechnung der Länder VGR dL - (National Accounts at the level of the Länder). Website.
URL `www.vgrdl.de`

Steinmann, L., Telser, H., Zweifel, P., 2007. Aging and future healthcare expenditure: A consistent approach. Forum for Health Economics and Policy 10 (2), Article 1.

Street, D., Bunch, D., Moore, B., 2001. Optimal designs for 2^k paired comparison experiments. Communications in Statistics: Theory and Methods 30, 2149–2171.

SwissDRG, 2009. Swissdrg. Website.
URL `www.swissdrg.org`

Telser, H., Vaterlaus, S., Zweifel, P., Eugster, P., 2004. Was leistet unser Gesundheitswesen? - (What Does Our Health System Achieve?). Rüegger Verlag, Zürich.

Telser, H., Zweifel, P., 2002. Measuring willingness-to-pay for risk reduction: An application of conjoint analysis. Health Economics 11, 129–139.

UK Prospective Diabetes Study (UKPDS) Group, 1998. Intensive blood-glucose control with sulphonylureas or insulin compared with conventional treatment and risk of complications in patients with type 2 diabetes (UKPDS 33). Lancet 352, 837–853.

Umpierrez, G., Hor, T., Smiley, D., Temponi, A., Umpierrez, D., Ceron, M., Munoz, C., Peng, L., Baldwin, D., 2009. Comparison of inpatient insulin regimes with detemir plus aspart versus neutral protamine Hagedorn plus regular in medical patients with type 2 diabetes. The Journal of Clinical Endocrinology and Metabolism 94 (2), 564–569.

Vague, P., Selam, J., Skeie, S., 2002. Insulin detemir is associated with more predictable glycemic control and lower risk of hypoglycemia compared to NPH insulin in subjects with type 1 diabetes on a basal-bolus regimen with premeal insulin aspart. Diabetes Care 26, 590–596.

Valentine, W. J., Erny-Albrecht, K., Ray, J., Roze, S., Cobden, D., Palmer, A., 2007. Therapy conversion to insulin detemir among patients with type 2 diabetes treated with oral agents: A modeling study of cost-effectiveness in the United States. Current Medical Research and Opinion 23 (4), 895–901(7).

Valentine, W. J., Palmer, A. J., Erny-Albrecht, K. M., Ray, J. A., Cobden, D., Foos, V., Lurati, F. M., Roze, S., 2006. Cost-effectiveness of basal insulin from a US health

system perspective: Comparative analyses of detemir, glargine, and NPH. Advances in Therapy 23 (2), 191–207.

Van Barneveld, E., Lamers, L., van Vliet, R., van de Ven, W., 2001. Risk sharing as a supplement to imperfect capitation: A tradeoff between selection and efficiency. Journal of Health Economics 20 (2), 147 – 168.

Van de Ven, W., Beck, K., Buchner, F., Chernikovsky, D., Gardiol, L., Holly, A., Lamers, L., Schokkaert, E., Shmueli, A., Spycher, S., Can de Voorde, C., van Vliet, R. C. J. A., Wasem, J., Zmora, I., 2003. Risk adjustment and risk selection on the sickness fund insurance market in five European countries. Health Policy 65, 75–98.

Van de Ven, W., Beck, K., Van de Voorde, C., Wasem, J., Zmora, I., 2007. Risk adjustment and risk selection in Europe: Six years later. Health Policy 83, 162–179.

Van de Ven, W., Van Vliet, R., Lamers, L. M., 2004. Health-adjusted premium subsidies in the Netherlands. Health Affairs 23 (3), 45–55.

Van de Ven, W. P., Schut, F. T., 2008. Universal mandatory health insurance in the Netherlands: A model for the United States? Health Affairs 27 (3), 771–781.

Wabitsch, M., Hauner, H., Hertrampf, M., Muche, R., Hay, B., Mayer, H., Debatin, K.-M., Heinze, E., 2004. Type 2 diabetes mellitus and impaired glucose regulation in obese German children and adolescents. International Journal of Obesity 28, 303–313.

Wild, S., Roglic, G., Green, A., Sicref, R., King, H., 2004. Global prevalence of diabetes: Estimates for the year 2000 and projections for 2030. Diabetes Care 27, 1047–1053.

Willen, A. R., O'Brien, B., 1996. Confidence intervals for cost-effectiveness ratios: An application of Fieller's theorem. Health Economics 5, 297–305.

Williams, A., Cookson, R., 2000. Equity in health. In: Culyer, A., Newhouse, J. (Eds.), Handbook of Health Economics. Vol. 1B. Elsevier, Amsterdam, pp. 1863–1910.

World Health Organization, 2007. Diabetes facts. Website.
URL `www.who.int/dietphysicalactivity/publications/facts/diabetes/en/`

World Health Organization, 2010. BMI classification. Website.
URL `http://apps.who.int/bmi/index.jsp?introPage=intro_3.html`

Zweifel, P., 2007. The theory of social health insurance. Foundations and Trends in Microeconomics 3 (3), 183–273.

Zweifel, P., Breuer, M., 2006. The case for risk-based premiums in public health insurance. Health Economics, Policy and Law 1 (2), 171–188.

Zweifel, P., Breyer, F., Kifmann, M., 2009. Health Economics, 2nd Edition. Springer Verlag, Berlin.

Zweifel, P., Eisen, R., 2005. Versicherungsökonomie - (Insurance Economics), 2nd Edition. Springer Verlag, Heidelberg.

Zweifel, P., Telser, H., Vaterlaus, S., 2006. Consumer resistance against regulation: The case of health care. Journal of Regulatory Economics 29 (3), 319–332.

Curriculum Vitae

Michèle Sennhauser was born on December 10, 1983 in Lucerne (Switzerland). She studied economics at the University of Zurich from 2002 to 2006 and graduated with a diploma thesis in econometrics. From 2006 to 2011 she has been a research assistant at the chair of Prof. Dr. Peter Zweifel (Department of Economics, University of Zurich) and wrote her doctoral thesis in health and insurance economics. One year of her doctoral program she studied at the chair of Prof. Willard G. Manning (The Harris School of Public Policy Studies, University of Chicago, USA). Since 2006 she has worked as a consultant for various health insurers and for the pharmaceutical industry.